



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

도시계획학 석사학위논문

로짓계열 수단선택모형의
추정 및 적합성 비교 분석

A Comparative Analysis of
Logit Mode Choice Models Examining
Estimation and Suitability

2018년 2월

서울대학교 환경대학원

환경계획학과 교통학전공

김 민 주

국문초록

기술의 발전과 사회가 변화함에 따라 과거에 비해 개인의 특성이 다양해지고, 이용 가능한 교통수단의 종류가 증가하였다. 이러한 다양성으로 인해 기존에 수단선택 로짓모형으로 사용되던 이항로짓모형이나 다항로짓모형 이외의 개인의 다양성을 반영한 발전된 모형의 필요성이 대두되고 있다.

개인의 선호다양성을 고려한 수단선택 로짓모형에 관한 연구는 국내외적으로 다양하게 진행되고 있지만, 그 역사는 그리 길지 않으며 본 연구에서 다루고자 하는 다항로짓모형, 혼합로짓모형, 잠재계층모형 세 모형의 비교연구는 아직까지 이루어지지 않았다.

대부분의 선행 연구는 특정 상황에 대한 설문조사를 기반으로 한 연구이다. 설문조사를 기반으로 할 경우 그 상황에서 응답자의 실제 행동과는 차이가 발생할 수 있다는 문제점이 있다. 따라서 본 연구에서는 서울시에서 개인의 실제 통행을 기반으로 하여 5년마다 실시하는 가구통행실태조사 자료를 이용하여, 서울시 통근자를 대상으로 개인의 선호다양성을 고려한 수단선택 로짓모형을 추정하고자 한다.

수단선택 모형 추정 시 이용하는 변수에는 자료에서 관측 가능한 통행시간, 통행비용 변수뿐만 아니라 사회·경제적 특성변수를 포함시켜, 개인의 다양성을 고려하였고, 소득에 따라 동일한 통행요금이라도 개인이 체감하는 비용은 다르기 때문에 월평균 분당소득 대비 통행비용을 이용하였다. 이용자가 이용할 수 있는 수단은 승용승합, 승용차동승, 버스, 지하철 4가지로 한정하였고, 승용차동승을 고려하는 점에서 공유시장이 점차 확대되고 있는 현재의 모습을 반영하였다는 점에서 선행연구와의 차별성을 보인다.

개인의 사회·경제적 특성에 따라 교통수단에 대한 개인의 효용이 다를 것을 고려하여, 수단 별로 더미변수를 다르게 하여 효용함수 식을 산정하고 모형을 추정하였다. 추정된 모형의 적합성을 확인하기 위해 ρ^2 값

을 비교하였고, 그 결과 잠재계층모형의 ρ^2 값이 0.2986으로 가장 높게 나타나 설명력이 가장 높은 것으로 확인되었다. 또한 추정된 모형이 다른 표본에도 동일하게 적용이 되는지를 확인하기 위해 추정에 사용된 표본 외의 검증 표본을 이용하여 확인하는 과정을 거쳤다. 그 결과 모든 모형에서 50%이상의 통행자가 실제로 효율이 가장 높은 수단을 이용한 것으로 나타났으며, 실제분담률과 추정분담률을 비교하여, 잠재계층모형에서 가장 비슷한 수단분담률을 보이는 것을 확인하였다.

또한 어떤 모형이 가장 개인의 선호다양성을 나타내는지 확인하기 위해 통행시간 한 단위당 개인의 통행시간가치와 통행시간과 통행비용의 증감에 따른 수요의 탄력성을 구하였다. 그 결과, 통행시간가치는 모형에 따라 다르게 나타났으며, 혼합로짓모형의 경우 절단된 정규분포보다 절단된 삼각분포에서 통행시간가치가 더 크게 나타났다. 삼각분포는 정규분포에 비해 극한값을 덜 가지며 값이 골고루 퍼져있어 상대적으로 큰 값이 더 많이 분포하여 통행시간가치가 높게 나타난 것으로 보인다. 잠재계층모형의 경우에는 계층에 속하는 개인의 특성에 따라 통행시간가치가 다르게 나타나 어떠한 기준으로 계층을 구분하느냐에 따라 그 값이 달라짐을 알 수 있었다.

통행시간과 통행비용의 수요탄력성은 모두 잠재계층모형일 때 가장 민감한 것으로 나타나, 잠재계층모형이 가장 개인의 선호다양성을 잘 반영하는 모형임을 알 수 있었다. 통행시간과 통행비용의 수단별 탄력성을 비교해보면, 모든 모형에서 지하철의 탄력성이 가장 덜 민감한 것으로 나타나고, 대부분의 모형과 수단에서 시간에 대한 탄력성이 비용에 대한 탄력성 보다 더 크게 나타났다. 이는 오전 첨두시의 통근자만을 대상으로 했기 때문에 통근통행의 특성 상 정시성이 수단선택에 중요한 결정요인으로 시간의 변화에 통근자가 더 민감하고, 지하철의 통행시간이나 통행비용이 증가하더라도 다른 수단에 비해 정시성이 보장되는 수단이므로 수요가 덜 감소하는 것으로 나타났다.

본 연구는 수단선택 로짓모형을 개인의 선호다양성을 고려한 다양한 모형으로 추정하고 이를 해석하여 잠재계층모형이 가장 설명력이 높은 모형임을 도출함으로써, 점차 세분화되는 개인의 특성을 잘 반영할 수

있는 모형을 제안하는 것에 의의가 있다. 따라서 현재 이용하고 있는 일반적인 수단선택 로짓모형 대신 개인의 선호다양성을 반영한 모형으로 개편이 필요하다.

주요어 : 수단선택모형, 다항로짓모형, 혼합로짓모형, 잠재계층모형, 선호다양성, 통행시간가치

학 번 : 2016-24830

목 차

제 1 장 서론	1
제 1 절 연구의 배경 및 목적	1
제 2 절 연구의 범위	3
제 3 절 연구의 절차	3
 제 2 장 모형 및 선행연구의 고찰	6
제 1 절 수단선택 로짓모형	6
1. 확률효용이론	6
2. 다항로짓모형	7
3. 선호다양성을 고려한 로짓모형	8
(1) 선호다양성	8
(2) 혼합로짓모형	8
(3) 잠재계층모형	9
제 2 절 선호다양성을 고려한 로짓모형을 추정한 선행 연구	12
1. 하나의 모형을 추정한 연구	12
2. 둘 이상의 모형을 추정한 연구	14
제 3 절 시사점	21
 제 3 장 연구방법론의 정립	22
제 1 절 수단선택 로짓모형의 설정	22

제 2 절 로짓모형 간 비교 분석 방법	27
1. 모형의 적합성	27
2. 모형의 추정결과	30
 제 4 장 자료의 구축	 33
제 1 절 서울시의 현황 분석	33
1. 서울시	33
2. 서울시 통근자	35
제 2 절 자료의 구축 및 특성 분석	37
1. 자료의 구축	37
(1) 종속변수	37
(2) 통행시간, 통행비용	39
2. 자료의 특성 분석	42
 제 5 장 모형의 추정 결과	 48
제 1 절 수단선택 로짓모형	48
1. 다항로짓모형	48
2. 혼합로짓모형	49
3. 잠재계층모형	51
제 2 절 모형 간 추정결과의 비교 분석	57
1. 적합성	57
(1) 계수의 통계적 유의성	57
(2) 적용 및 확인	68
2. 시간가치 및 탄력성	60

제 6 장 결론	64
제 1 절 결론	64
제 2 절 한계 및 향후 연구과제	66
참고문헌	68
부록	72
<부록 1-1> 수단별 개인의 사회·경제적 특성(추정)	72
<부록 1-2> 개인의 사회·경제적 특성별 수단분담률(추정)	73
<부록 1-3> 수단별 개인의 사회·경제적 특성 비율(추정)	74
<부록 2-1> 수단별 개인의 사회·경제적 특성(적용)	75
<부록 2-2> 개인의 사회·경제적 특성별 수단분담률(적용)	76
<부록 2-3> 수단별 개인의 사회·경제적 특성 비율(적용)	77
<부록 3-1> 승용차보유가구 통행시간분포	78
<부록 3-2> 승용차보유가구 통행비용분포	78
<부록 3-3> 소득대비 통행비용 분포	78
<부록 3-4> 로그통행시간 분포	79
<부록 3-5> 로그통행비용 분포	79
<부록 3-6> 로그 소득대비 통행비용분포	79
Abstract	80

표 목 차

<표 1> 연구의 범위	3
<표 2> 수단선택 로짓모형별 특징 및 한계점 비교	11
<표 3> 정진혁 외(2007) 수단선택 로짓모형의 계수 분포	14
<표 4> 선호다양성을 고려한 수단선택 로짓모형 간 비교에 관한 연구	19
<표 5> 개인의 선호다양성을 고려한 수단선택 로짓모형 선행연구 고찰	20
<표 6> 소득계층 별 적용 분당 소득	22
<표 7> 변수 설명 및 표기 방법	23
<표 8> 개인의 사회·경제적 속성 더미변수 설정	24
<표 9> 서울시 통행목적별 통행량	33
<표 10> 서울시 통행수단별 통행량	34
<표 11> 가구당 자동차 보유대수	34
<표 12> 서울시 가구주 연령	34
<표 13> 서울시 내 통근통행 수단분담률	35
<표 14> 서울시 통근 가구주 연령	36
<표 15> 승용차 보유/미보유 가구별 통근 통행량	37
<표 16> 승용차 보유 가구별 통근통행 수단분담률	38
<표 17> 추정과 검증에 사용되는 표본의 수단분담률	39
<표 18> 속도별 승용차 운영비용	41
<표 19> 서울시 평균 출근 지표	41
<표 20> 표본의 사회·경제적 특성	43
<표 21> 표본의 통행시간과 통행비용 분포	46

<표 22> 수단별 통행시간과 통행비용 별 Quartile	46
<표 23> 통행시간과 통행비용 변수의 분포	47
<표 24> 다항로짓모형 추정 결과	49
<표 25> 혼합로짓모형 추정 결과	50
<표 26> 최적 계층 수 결정 기준	52
<표 27> 잠재계층모형 계층 구분 및 계층 특성	52
<표 28> 잠재계층모형의 계층 별 수단분담률	53
<표 29> 잠재계층모형 추정 결과	54
<표 30> 수단선택 로짓모형 추정 결과	56
<표 31> 모형의 적합성 비교	57
<표 32> 최대효용수단과 실제이용수단의 일치 여부	58
<표 33> 적용 표본의 실제분담률과 추정분담률	59
<표 34> 통행시간가치	60
<표 35> 선행연구 중 시간가치를 비교한 연구	61
<표 36> 로짓모형 별 통행시간과 통행비용에 대한 수요 탄력성	62

그 립 목 차

<그림 1> 연구흐름도	5
<그림 2> 통근통행 출발시간 분포	36
<그림 3> 네이버 지도 검색 예시	40
<그림 4> 표본의 사회·경제적 특성(추정)	44
<그림 5> 표본의 사회·경제적 특성(적용)	44

제 1 장 서론

제 1 절 연구의 배경 및 목적

통행발생, 통행분포, 수단분담, 통행배분의 4단계 수요추정방법은 교통계획에서 오래전부터 교통수요를 예측하는 방법으로 이용되어 왔다. 도로의 건설이나 신규 지하철 노선 계획과 같은 새로운 사업을 수립하기 전, 예비타당성 조사를 실시할 때에도 위의 방법을 토대로 미래의 교통수요를 예측하고 사업의 타당성을 분석한다.

이처럼 4단계 수요추정방법은 교통 분야를 계획하고 정책을 시행할 때 필수적이다. 개별 교통수단에 대한 수요의 예측이 중요하지만, 교통수요자 개개인의 통행은 너무 다양해서 예측하기 힘들다. 따라서 지금까지는 모든 개인에게 수단이 미치는 영향이 동일하다고 가정하고 수단선택 로짓모형을 추정하였으며, 수단선택 로짓모형으로는 로짓모형(Logit Model)을 주로 사용하며, 대안의 수에 따라 이항로짓모형(Binary Logit Model)과 다항로짓모형(Multinomial Logit Model)을 이용했다.

이항로짓모형과 다항로짓모형은 기본적으로 개인의 선호다양성을 고려하지 않으며 IIA(Independence of Irrelevant Alternatives)를 가정한다는 문제점이 있다.¹⁾ 이러한 모형의 단점을 극복하기 위해 발전된 모형을 제시하는 연구가 다양하게 이루어지고 있으며, 대표적으로 혼합로짓모형(Mixed Logit Model), 잠재계층모형(Latent Class Model), 다항프로빗모형(Multinomial Probit Model) 등이 있다. 선행연구에서 개인의 선호다양성을 반영한 모형이 일반적인 로짓모형보다 더 적합하다고 결과가 나타났지만, 그 중에서도 어떤 모형이 정책을 수립하거나 계획을 할 때 가장

1) IIA는 여러 대안 가운데 한 대안을 선택할 확률이 다른 대안 존재 유무에 영향을 받지 않는다고 가정한다. (Louviere et al, 2000).

적절한 지에 대해서는 절대적인 지표가 없다. 하지만, 선호다양성을 반영한 모형이 개인의 실제 선택을 더 잘 반영한다는 것은 분명하다.

기술의 발전으로 인해 자율주행자동차(Autonomous Car), 스마트 모빌리티(Smart Mobility)와 같은 새로운 교통수단이 등장하고, 공유경제의 확산으로 카풀(Carpool)이나 카셰어링(Car-sharing)²⁾, 대체에너지를 이용하는 전기차(Plug-in Electric Vehicle)나 수소전기자동차(Hydrogen Electric Vehicle)와 같은 이전에는 없던 대안도 활성화되고 있다. 대표적으로 가장 널리 알려진 차량 공유경제 기업인 우버(Uber)의 경우, 2010년 시장 규모 8.5억 달러에서 5년 만에 100억 달러로 시장의 규모가 커졌으며³⁾, 우버 이후 다양한 카풀 회사들이 등장하였고 한국에서도 2016년부터 럭시(Luxi)와 풀러스(Poolus)와 같은 카풀 회사에서 서비스를 제공하고 있다.

이처럼 점차 개인이 선택할 수 있는 대안의 수가 늘어나고 각 대안의 특징이 명확하게 구분되면서, 이용자의 선호다양성 또한 더욱 다양해질 것이다. 그럼에도 불구하고 현재 사용하고 있는 수요예측모형으로는 개인의 선호다양성을 고려하지 못하고 있기 때문에, 더욱 다양해질 개인의 선호도를 반영하기 위한 발전된 수단선택 로짓모형이 제안되어야 할 것이다. 따라서 본 연구에서는 개인의 선호다양성을 고려하는 수단선택 로짓모형을 추정하여 일반적인 로짓모형과 비교하고, 추정된 계수의 유의성을 통해 어떤 모형이 개인의 다양성을 반영하기에 가장 적합한 지를 확인한다. 다른 연구와 달리 본 연구에서는 승용차를 승용승합과 승용자동차승으로 개인의 선택 대안을 구분하여 향후 새로운 교통수단에 대한 개인의 다양성이 확대되었을 때의 상황에 대해서도 고려한다.

2) 카풀(Carpool)은 Car-sharing, Ride-sharing, Life-sharing이라고도 하는데, 카셰어링(Car-sharing)은 일부 국가에서 하나의 자동차를 사람들이 돌려가면서 타는 것을 의미한다. 따라서 의미 혼동 가능성을 고려해 본 연구에서는 카풀은 승용차에 동승하는 개념으로 사용하고, 카셰어링은 공유차량을 본인이 직접 운전하는 개념으로 사용하도록 한다.

3) 자료는 미국 시장조사기관 massolution을 참고하였다. (www.massolution.com)

제 2 절 연구의 범위

본 연구의 공간적 범위는 서울특별시이며, 통행의 출·도착지를 수도권권을 제외한 서울시로 제한한다. 그 이유는 출근시간에 수도권에서 서울시로 유입하는 통행자들의 행태, 서울시에서 수도권으로 유출되는 통행자들의 행태, 서울시 내부 통행자들의 행태가 각기 다른 특징을 지니고 있기 때문이다. 또한, 수도권과 서울시 간의 통행은 장거리 통행인 경우가 많고, 연구의 시간적 범위가 수도권과 서울시를 잇는 주요 노선인 ‘신분당선’, ‘분당선’, ‘경춘선’ 등이 개통되지 않거나 현재의 노선과 같은 모습을 보이지 않던 기간으로 설정하였기 때문에 선택 대안의 수가 한정적일 것으로 예상되기 때문이다.

본 연구에서 이용하고자 하는 자료는 가구통행실태조사 자료로 공개된 가장 최신 자료는 2010년 자료이므로, 시간적 범위를 2010년으로 한정한다. 연구의 대상은 통근통행자로 선정하였는데, 이는 통근통행은 일반적인 통행보다 수단선택의 변동이 적고, 통행시간에 민감하기 때문에 다른 통행보다 뚜렷한 통행행태를 보일 것으로 예상되기 때문이다.

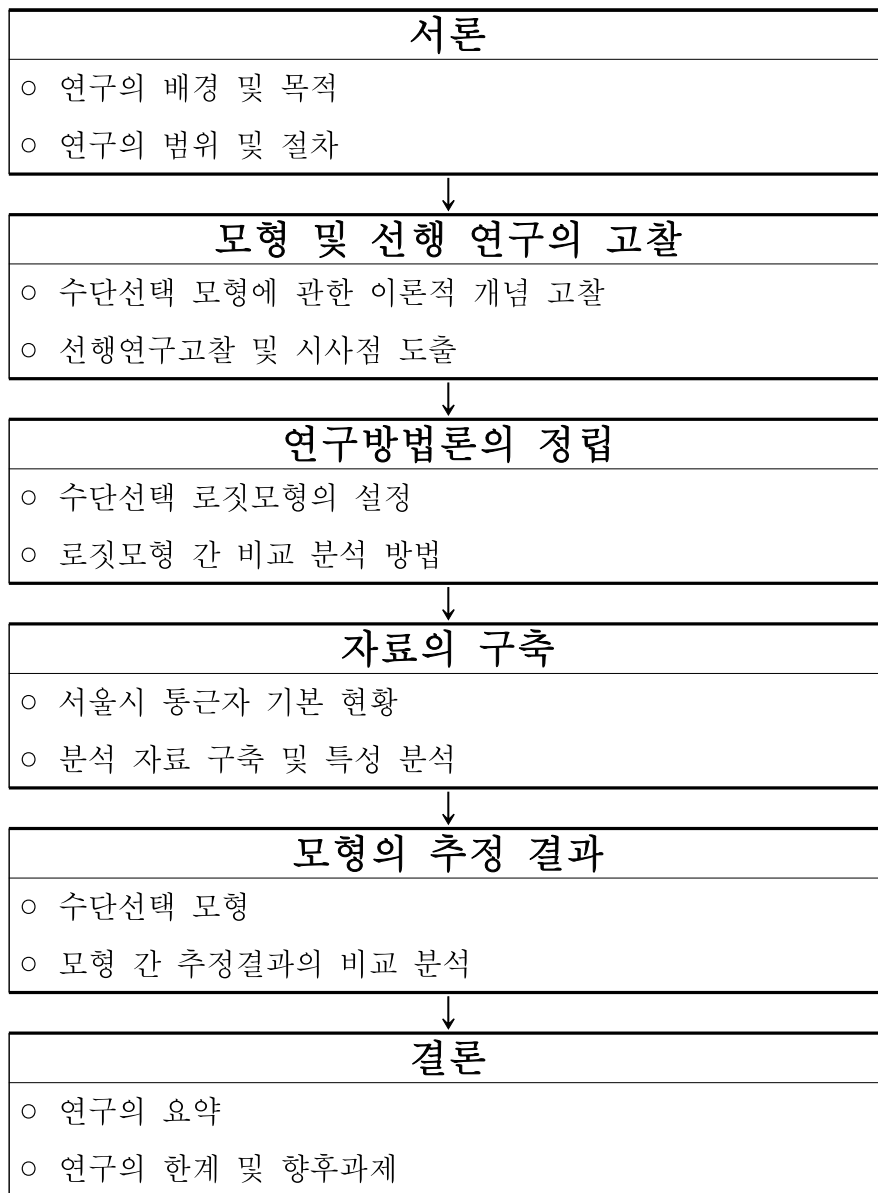
<표 1> 연구의 범위

공간적 범위	서울특별시
시간적 범위	2010년
대상	통근통행자
이용자료	가구통행실태조사자료

제 3 절 연구의 절차

본 연구는 보편적으로 사용하는 수단선택 모형인 다항로짓모형과 개인의 선호다양성을 고려하는 수단선택 모형을 추정하고 결과를 비교하는

데 목적이 있다. 따라서 본 연구는 각각의 수단선택 모형에 관한 이론적 개념에 대해 고찰하고, 특히 개인의 선호다양성을 고려하는 모형(혼합로짓모형, 잠재계층모형)을 추정한 선행연구들과 여러 가지 모형 간의 비교연구를 한 선행연구들을 살펴본다. 그리고 개인의 다양한 특성을 고려한 수단선택 로짓모형을 설정하고, 모형 간 비교분석 방법에 대해 살펴본다. 그 다음 본 연구에서 이용하는 기본 자료인 가구통행실태조사 자료에 대한 기술적 분석과 분석에 이용하는 표본을 구축한다. 구축된 표본을 이용하여 각각의 수단선택 모형을 추정하고, 이를 토대로 어떤 모형이 개인의 선호다양성을 반영한 수단선택 모형으로 적절한지에 대해 모형 간 비교를 진행할 것이다. 또한 추정된 식이 제대로 추정됐는지 확인하기 위해 추정에는 사용되지 않은 자료를 이용하여, 효용함수 식이 얼마나 잘 추정되었는지 적용시켜본다.



<그림 1> 연구흐름도

제 2 장 모형 및 선행연구의 고찰

제 1 절 수단선택 로짓모형

1) 확률효용이론

로짓모형(Logit Model)은 확률효용이론(Random Utility Theory)을 기반으로 한 이산선택모형(Discrete Choice Model)으로, 모든 사람은 주어진 대안 중에서 가장 큰 효용을 얻을 수 있는 대안을 선택한다는 가정을 전제로 한다. 확률효용이론에서 대안 별 총 효용(Total Utility)은 결정적 효용(Deterministic Utility)과 확률적 효용(Random Utility)으로 구분되며, 이를 식으로 나타내면 식 (1)과 같다.

$$U_{nj} = V_{nj} + \varepsilon_{nj} \quad \dots(1)$$

U_{nj} 는 개인 n 의 대안 j 에 대한 총효용을 나타내며, V_{nj} 는 개인 n 의 대안 j 에 대한 결정적 효용, ε_{nj} 는 개인 n 의 대안 j 에 대한 확률적 효용이다. 결정적 효용과 확률적 효용은 각각 관측이 가능한 요소와 관측할 수 없는 요소를 말하며,⁴⁾ 결정적 효용은 성별, 연령, 소득, 직업종류 등과 같은 개인의 사회·경제적 특성과 통행시간, 통행비용, 통행거리 등과 같은 특정 대안의 특성으로 구성되고, 확률적 효용은 자료로부터 도출할 수 없는 특성, 불완전한 정보 등을 의미한다.

이러한 이론을 토대로 본 연구에서는 개인의 효용을 극대화하는 수단선택 모형을 다양한 방법으로 추정한다.

4) 윤대식(2001), 「교통수요분석」, 서울: 박영사.

2) 다항로짓모형

수단선택 모형 추정 시, 3개 이상의 대안이 있는 경우에는 다항로짓 모형(Multinomial Logit Model)을 사용한다. 다항로짓모형은 선택 가능한 대안이 다항분포(Multinomial Distribution)를 이루고 있는 경우로 이항선택모형(Binomial Distribution)과 차이를 보인다. 수단선택 시 사용하는 로짓모형은 성별, 연령, 소득, 직업종류 등과 같이 대안에 따라 변화하지 않는 개인의 사회·경제적인 특성뿐만 아니라 선택대안에 따라 변하는 통행시간, 통행비용, 통행거리 등과 같은 변수까지 고려할 수 있다.⁵⁾

따라서 개인 n 에 대한 대안 j 에 관한 효용함수는 식 (2)와 같이 나타나고, X_{nj} 는 개인의 특성이나 대안의 특성과 같은 설명변수를 나타낸다. 이 모형의 경우에는 각각의 대안이 개인에게 미치는 계수 값이 β' 로 동일하게 추정된다. 여기서 개인 n 이 대안 j 를 선택할 확률 $P_n(j)$ 는 식 (3)과 같이 나타낼 수 있다.

$$U_n(j) = \theta_{nj} + \beta' X_{nj} + \varepsilon_{nj} \quad \dots(2)$$

$$P_n(j) = \frac{\exp(\theta_{nj} + \beta' X_{nj})}{\sum_{k=1}^J \exp(\theta_{nk} + \beta' X_{nk})} \quad \dots(3)$$

다항로짓모형은 수단선택 모형 중에서 다른 모형에 비해 비교적 계산이 용이하다. 하지만, 개인의 선호다양성을 고려하지 못하고, IIA 가정을 수반한다는 한계가 있다.

5) 이희연 외(2012), 「고급통계분석론 - 이론과 실습」, 경기: 문우사.

3) 선호다양성을 고려한 로짓모형

(1) 선호다양성

선호다양성(Preference Variation)은 동일한 상황에 대해서도 개인의 사회·경제적 특성뿐만 아니라 개인이 처한 상황에 따라 개인에게 미치는 영향이 다르게 나타나는 것을 의미한다.

기술이 발전하고 사회가 변하면서 개인의 특성은 점점 세분화되고, 이로 인해 개인의 선호다양성 또한 더욱 분명하게 나타난다. 따라서 계획이나 정책을 수립할 때, 모든 사람을 동일하다고 가정하고 계획하는 것에 대한 한계와 개인의 선호다양성을 반영하는 것에 대한 필요성이 제기되면서, 선호다양성을 고려한 다양한 형태의 로짓모형이 등장하였다.

(2) 혼합로짓모형

혼합로짓모형(Mixed Logit Model)은 일반 로짓모형의 한계점을 해결하기 위해 발전된 대표적인 모형으로, 설명 변수의 계수 추정 값이 개인별로 다르며 계수는 특정 확률분포를 따른다고 본다.

혼합로짓모형은 일반 로짓모형과 다르게 식 (4)에서처럼 변수의 계수 값을 연구자가 정규분포(Normal Distribution), 로그정규분포(Log-normal Distribution), 삼각분포(Triangular Distribution), 균등분포(Uniform Distribution)등의 확률밀도함수로 계수 분포를 설정할 수 있으며, 식 (5)와 같이 일반 로짓모형에 확률밀도함수 $f(\beta)$ 를 곱하는 형태로 표현된다.

$$U_n(j) = \theta_{nj} + \beta'_n X_{nj} + \varepsilon_{nj} \quad \dots(4)$$

$$P_n(j) = \int \left(\frac{\exp(\theta_{nj} + \beta' X_{nj})}{\sum_{k=1}^J \exp(\theta_k + \beta' X_{nk})} \right) f(\beta) d\beta \quad \dots(5)$$

혼합로짓모형은 변수별 계수를 확률분포로 가정함으로써 탄력적으로 개인의 선호다양성을 표현할 수 있고, IIA에 대한 가정을 완화시킬 수 있다는 점에서 일반 로짓모형보다 발전된 형태이다. 하지만, 미지의 확률 분포를 가정해야한다는 점과 연구자가 가정하는 확률분포에 따라 결과 값이 다르게 나타난다는 문제가 있다. 따라서 개인의 특성과 관련이 있는 선호다양성의 근본적인 원인을 확인하는 데에는 적합하지 않다.⁶⁾

(3) 잠재계층모형

특정한 계층을 대상으로 하는 마케팅 분야에서 주로 사용되는 잠재 계층모형(Latent Class Model)은 비슷한 특성을 띄는 2개 이상의 계층으로 표본을 구분하여 각 계층의 모수를 각각 추정하는 방법이다. 잠재계층모형은 계층을 나눈다는 점에서 군집분석(Cluster Analysis)과 비슷하지만, 모형을 기반으로 하여 다양한 종류의 자료 분석에 이용가능하고, 통계적 기준을 가지고 있다는 점에서 군집분석과 차이가 있다.⁷⁾

잠재계층모형에서는 하나의 잠재적인 계층에 개인이 속하게 되며, 계층 내 선호는 동질적이지만 계층 간에는 선호다양성이 존재한다고 가정한다. 계층 s 에 속한 개인 n 이 대안 j 를 선택할 확률은 식 (7)과 같이 나타난다.

6) Boxwall et al.(1999), *Understanding Heterogeneous Preferences in Random Utility Models: The Use of Latent Class Analysis*, University of Alberta Edmonton.

7) 배윤경 외(2011)에서 인용하였다.

$$U_n(j) = \theta_{sj} + \beta'_s X_{nj} + \varepsilon_{nj} \quad \dots(6)$$

$$P_n(j) = \frac{\exp(\theta_{sj} + \beta'_s X_{nj})}{\sum_{k=1}^J \exp(\theta_{sk} + \beta'_s X_{nk})} \quad \dots(7)$$

잠재계층모형에서 최적의 계층 수를 결정하는 방법으로는 *BIC*(the Bayesian Information Criterion), *AIC*(the Akaike Information Criterion) 이 있고 보통은 이들의 최솟값으로 계층의 수를 결정한다. *BIC*와 *AIC* 를 구하는 식은 각각 식 (8)과 식 (9)이고, $LL(\beta)$ 는 로그가능도⁸⁾ 값, p 는 계수 추정치의 수, N 은 응답자 수를 나타낸다.

$$BIC = -2 \cdot LL(\beta) + p \cdot \log(N) \quad \dots(8)$$

$$AIC = -2 \cdot LL(\beta) + p \quad \dots(9)$$

잠재계층모형은 혼합로지트모형에 비해 선호다양성의 근본적인 원인규명에 용이하지만, 동일한 계층 내 선호가 동질적이라는 가정이 너무 제한적이고 이산값 사이의 모수 누락의 문제가 있다.⁹⁾

본 연구에서는 앞서 살펴본 세 가지 모형을 각각 추정하여, 각 모형 별로 결과가 어떻게 나타나는지 서로 비교해본다. 이미 여러 연구에서 모형 간 비교연구가 진행되었지만, 어느 모형이 절대적으로 우월하다는 연구는 나타나지 않는다. 이와 관련된 선행연구 및 분석 방법, 모형 간 차이에 대해서는 다음 절에서 자세히 살펴보도록 하겠다.

8) 보통 Likelihood를 우도라고 많이 사용하지만, 이는 일본어로 해석한 용어이므로 본 연구에서는 우도대신 가능도로 대체하겠다.

9) Bujosa et al.(2010)를 참고하였다.

<표 2> 수단선택 로짓모형별 특징 및 한계점 비교

	다항로짓모형	혼합로짓모형	잠재계층모형
식	$U_n(j) = \theta_{nj} + \beta' X_{nj} + \varepsilon_{nj}$	$U_n(j) = \theta_{nj} + \beta'_n X_{nj} + \varepsilon_{nj}$	$U_n(j) = \theta_{sj} + \beta'_s X_{nj} + \varepsilon_{nj}$
기준		- 연구자가 임의로 계수의 확률분포를 가정	- 계층의 수는 <i>BIC</i> 와 <i>AIC</i> 의 최솟값으로 결정
특징	- 3가지 이상의 대안 가운데 개인의 효용이 극대화되는 대안을 선택함	- 계수의 확률분포를 연구자가 가정할 수 있어 탄력적임 - 개인의 선호다양성을 고려함 - IIA 가정을 완화시킴	- 동질성을 띄는 개인을 하나의 계층으로 구분하고, 각 계층의 계수를 각각 추정함 - 선호다양성의 원인규명에 적합함 - 계층 간 선호다양성을 고려함 - 사전 확률분포를 가정하지 않아도 됨
한계	- 개인의 선호 다양성을 고려하지 않음 - IIA 가정	- 미지의 모수 확률분포를 가정함 - 선호다양성의 원인규명에 적합하지 않음	- 계층 내의 선호 다양성을 고려하지 않음 - 이산 값 사이의 모수 누락의 문제가 있음

제 2 절 선호다양성을 고려한 로짓모형을 추정한 선행연구

1) 하나의 모형을 추정한 연구

선호다양성을 고려한 연구의 필요성이 제기되면서 교통 분야에서도 연구가 진행되고는 있지만, 국내에서는 아직까지 일반 로짓모형에 비해 많이 연구되지 않은 상황이다. 배운경 외(2010)는 개인의 심리적인 요인이나 선호도가 수단선택에 영향을 미친다고 보고, 한강 수상교통이 도입됐을 때의 수단선택 모형을 추정하였다. 전체 표본을 대상으로 한 수단선택 로짓모형과 비교해보았을 때, 잠재계층으로 분리한 수단선택 로짓모형의 설명력이 더 높아졌고, 이를 통해 개인이나 대안의 특성변수 뿐만 아니라 개인의 잠재적인 선호도가 미치는 영향 또한 중요하게 고려되어야 할 부분임을 강조하였다.

배운경 외(2011)의 다른 연구에서는 개인이나 수단의 특성 외에도 주변 환경이나 토지이용특성이 수단선택에 많은 영향을 미친다는 것을 확인하였다. *BIC*와 *AIC*기준에 의해 아파트집중지역 중 종사자수가 적은 지역, 대중교통 공급이 낙후된 지역, 종사자수와 인구밀도가 많은 지역, 아파트수와 지하철역수가 적지만 종사자수가 많은 지역, 총 4개의 계층으로 분류했다. 결과적으로 잠재계층별로 다른 토지이용특성과 수단선택행태를 보이는 것으로 나타났다.

추상호(2012)는 활동유형, 장소, 시간, 활동지속시간, 동반자 등의 활동요소가 개인의 수단선택에 영향을 미칠 것이라고 보고, 2006년 서울시 가구통행실태조사 자료를 이용하여 혼합로짓모형을 추정하였고, 그 때 통행시간 및 통행비용 계수 분포를 정규분포로 가정하였다. 결과적으로 쇼핑과 여가활동의 경우 통행비용에 대해서만 수단선택의 선호 다양성이 통계적으로 유의하게 나타났고, 기타활동의 경우에는 통행시간과 통행비

용의 계수 값이 분포가 아닌 고정된 값으로 추정되었다.

국외연구에서 Hensher and Greene(2001)는 정규분포, 로그정규분포, 삼각분포, 균등분포일 때의 장단점을 살펴보았다. 그 중에서 통행시간가치의 표준편차는 다른 분포에 비해 크게 나타나지만, 양의 값만 나와 통행시간의 변화가 모든 사람들에게 같은 방향으로 영향을 미칠 것이라는 일반적인 직관과 부합하는 로그정규분포로 계수 분포를 가정하였다. 계수는 통행시간만 분포로 가정할 때, 통행비용만 분포로 가정할 때, 통행시간과 통행비용 모두 분포로 가정할 때로 구분하여서 모형을 추정하였으며, 계수를 어떤 분포로, 분포의 표준편차를 어떻게 설정하느냐에 따라 통행시간가치가 달라지는 것을 확인하였다.

Atasoy et al.(2011)는 스위스를 대상으로 설문조사와 GPS장치에 의해 자료를 수집하였으며, 수집된 자료를 군집분석을 통해 계층을 구분하였다. 각각의 계층에 대해 수단선택 로짓모형을 추정하고, 추정된 모형의 계수 값을 이용하여 수요탄력성과 시간가치를 구하였다. 가족과 아이들과 함께 사는 고소득계층의 중년층(계층 1)과 혼자 살거나 자녀가 없는 부부(계층 2)로 계층을 구분하였으며, 계층별 수단선택 로짓모형의 수요탄력성과 시간가치를 산정하였다. 계층 2가 계층 1에 비해 개인수단과 대중교통 모두에서 통행시간과 통행비용에 더 민감한 것으로 나타났고, 계층 1의 시간가치가 개인수단과 대중교통 모두에서 더 높은 것으로 나타났다.

독일 두 개의 도시에서 6주간 통행실태조사를 한 Vij et al.(2013)의 연구는 승용차 운전자와 승용차 동승자 간의 특징은 다르게 나타날 것으로 기대되지만, 두 이용 수단은 독립적이지 않다고 보았다. 승용차 운전자 계층과 시간에 민감한 다양한 수단 이용자 계층, 시간에 민감하지 않은 다양한 수단 이용자 계층, 총 3가지 계층으로 구분하였다. 또한 각 계층별로 업무통행과 비업무 통행 시의 수단선택 로짓모형을 추정하고, 각 계층 별 통행시간에 대한 수요탄력성을 확인하였다. 계층을 구분하여 수

단선택 로짓모형을 추정하는 것은 향후 계층의 특성에 맞는 특정 정책을 시행하는데 더 유용할 것으로 예측되며, 이러한 방법은 모형의 예측력을 더 향상 시킬 것으로 나타났다.

2) 둘 이상의 모형을 추정한 연구

기존의 로짓모형의 한계점을 파악하고, 개인의 선호다양성을 고려한 모형과 일반 로짓모형 간 추정 결과를 비교한 연구는 국내외적으로 활발히 연구되고 있다. 하지만, 국내에서는 다항로짓모형과 혼합로짓모형간의 비교만 존재할 뿐, 본 연구에서 다루고자 하는 로짓모형, 혼합로짓모형, 잠재계층모형 전부를 비교하는 연구는 국외에서만 연구되고 있다.

국내에서는 정진혁 외(2007), 이준규 외(2015), 조신형 외(2017)의 논문에서 수단선택 로짓모형으로 다항로짓모형과 혼합로짓모형 간의 추정 결과를 비교하는 연구를 진행하였다. 정진혁 외(2007)에서는 2002년 서울시 가구통행실태조사 자료를 이용하여 다항로짓모형과 모수의 계수를 3가지 분포로 구분하여 혼합로짓모형을 추정하였다. 설정한 수단선택 로짓모형과 가정한 분포는 <표 3>과 같다.

<표 3> 정진혁 외(2007) 수단선택 로짓모형의 계수 분포

구분	MNL	ML-1	ML-2	ML-3
모형	로짓모형	혼합로짓모형	혼합로짓모형	혼합로짓모형
통행비용	고정계수	고정계수	고정계수	정규분포
통행시간	고정계수	균등분포	정규분포	정규분포

95% 신뢰구간에서의 통행시간가치는 혼합로짓모형과 다항로짓모형 간에 큰 차이를 보였고, 개인의 선호다양성을 고려하지 않을 경우에 통행시간가치가 과대 추정되거나 과소 추정 될 가능성이 있으므로 교통 계획 및 정책 평가 시 문제가 될 것이라 예상하였다.

이준규 외(2015)는 우리나라 내륙과 제주도 사이 해저터널로 연계한다는 가정 하에 서울-제주 노선 수단선택에 관한 연구를 진행하였으며, 운임과 통행소요시간 계수를 삼각분포로 가정하여 혼합로짓모형을 추정하였다. 비즈니스여객과 레저여객 모두 시간가치는 혼합로짓모형으로 추정한 경우에 더 높게 나타났다. 이를 통해 항공여객의 선호다양성을 확인 할 수 있었으며, 이러한 결과가 맞춤형 전략 수립, 항공사의 경쟁력 강화 등에 유용하게 활용될 것이라 기대하였다.

조신형 외(2017)는 통행수단에 대한 개인의 시간가치를 분석하고자 다항로짓모형과 개인의 선호다양성을 반영한 혼합로짓모형을 추정, 비교하였다. 분석 범위는 서울에서 부산까지이며, 선택 가능 대안인 승용차, 항공, 버스, 철도 4개의 수단에 대한 통행시간과 통행비용을 제시하고 수단에 대한 선호도를 설문 조사하였다. 혼합로짓모형의 경우에는 오차항의 분포를 정규분포와 삼각분포로 가정하였으며, 정규분포로 가정한 혼합로짓모형이 가장 설명력이 높은 것으로 나타났다. 또한 수단선택 로짓모형 별 통행시간 가치를 추정하여 한계임금율법과 비교하였다. 통행시간 가치는 통행 목적에 따라 다르게 나타났으며, 업무목적은 다항로짓모형의 통행시간가치가 가장 높았고, 비업무목적의 경우에는 삼각분포로 가정한 혼합로짓모형의 경우에 통행시간가치가 가장 높게 나타났다. 이처럼 관측되지 않은 개인의 선호도를 분포를 어떻게 가정하느냐에 따라 결과가 다르게 나타나는 것을 확인할 수 있었다.

Greene and Hensher(2003)은 뉴질랜드의 차량 운전자 자료를 이용하여 도로 환경에 대한 선호도 조사를 실시하였다. 혼합로짓모형의 통행시간 계수 분포를 삼각분포로 가정하였으며, 각각의 모형에서 대안 별 통행시간절감가치와 통행시간과 비용에 대한 탄력성을 비교하였다. 결과적으로 혼합로짓모형과 잠재계층모형 모두 관측되지 않은 개인의 선호다양성을 고려하는 모형으로, 혼합로짓모형은 완전한 계수를 이용하지만, 분포를 가정함으로써 개인의 관측되지 않은 선호다양성 내에서 충분히

유연하게 모형을 추정할 수 있었고, 잠재계층모형의 경우에는 개인의 선호다양성에 대해 강하거나 부적절한 분포 가정에서 벗어나 분석할 수 있는 장점을 가지고 있는 것을 확인하였다.

Lee et al.(2003)은 일본 히로시마의 새로운 교통수단과 자가용을 대상으로 혼합로짓모형과 잠재계층모형, 혼합-잠재계층모형(LC-MXL Model)을 비교하여 개인의 선호다양성을 가장 명확하게 나타낼 수 있는 모형을 제시하였다. 혼합-잠재계층모형은 계층을 먼저 나누고, 그 이후 모든 잠재계층에 대해 혼합로짓모형을 동시에 추정하여 계층 간 뿐만 아니라 계층 내의 선호 다양성 모두 고려하는 모형이다. 다항로짓모형, 혼합로짓모형, 잠재계층모형, 혼합-잠재계층모형으로 통행시간과 지체시간에 대한 정보, 교통 정보의 유무에 대해 시나리오를 설정하고 각각의 시나리오를 추정하였다. 그 결과, 혼합로짓모형은 교통 정보 제공에 대한 개인의 선호다양성을 효과적으로 반영할 수 있었으나, 개인의 선호다양성이 발생하는 이유에 대해서는 분석할 수 없었다. 반면에, 잠재계층모형의 경우에는 관측되지 않은 선호다양성을 분석하여 그러한 이유까지도 설명가능하며, 혼합-잠재계층모형의 Rho-squared값이 가장 높게 나타난 것으로 보아 계층 간과 계층 내 선호다양성 모두를 고려하는 경우에 모형의 설명력이 높아지는 것을 확인하였다.

Viton(2004)의 연구에서는 미국 샌프란시스코의 교통 시스템에서 수집된 자료를 이용하여 4개의 대안(승용차, 도보 접근으로 버스 이용, 차로 접근하여 버스 이용, 카풀)에 대한 수단선택 로짓모형을 추정하였다. Viton은 차내시간과 대기시간에 대해 정규분포, 로그정규분포, 균등분포, 삼각분포로 가정한 혼합로짓모형을 추정하였으나 결과적으로 균등분포와 삼각분포만 유의미한 결과가 나타났으며, 그 결과는 비슷하게 나타났다. 통행시간가치에서 차내시간가치는 다항로짓모형과 혼합로짓모형 두 모형이 비슷하게 나타났으나, 접근시간가치는 다항로짓모형에서, 대기시간가치는 혼합로짓모형에서 더 크게 나타났다.

Shen(2009)은 특히 시간절약가치, 직접선택탄력성, 선택예측 가능성에 대해 혼합로짓모형과 잠재계층모형의 차이를 분석하였다. 일본 오사카 시의 2개의 도시에 대한 설문자료를 이용하여 승용차, 버스, 전철에 대한 수단선택 로짓모형을 추정하였다. 혼합로짓모형의 경우 통행시간을 정규분포와 삼각분포로 가정하고 추정하였으나 두 개의 값이 비슷하게 추정되었다. 또한 두 도시의 모형별 통행시간절감가치와 탄력성을 추정하고 비교하였다. 결과적으로는 수요예측지수나 비내재적 선택모형에는 잠재계층모형이 혼합로짓모형보다 통계적으로 더 효과적으로 났지만, 잠재계층모형이 혼합로짓모형보다 절대적으로 우세한 모형은 아니므로 자료와 추정방법에 따라 적절한 모형을 사용하는 것이 중요하다고 결론지었다.

Greene and Hensher(2010)의 연구에서는 수단 선택에 있어 규모의 이질성(Scale Heterogeneity)와 선호의 이질성(Taste Heterogeneity) 사이의 관계를 확인하고자 하였다. 2003년 호주 시드니(Sydney)의 통근통행 자료를 이용하여 새로운 교통수단의 등장에 따른 수단선택 로짓모형을 추정하였다. 다항로짓모형(M1), 혼합로짓모형(M2), 규모와 선호의 이질성을 모두 고려한 일반화된 혼합로짓모형(Generalized Mixed Logit Model; M3), 규모의 이질성만 고려한 규모이질성 모형(Scale Heterogeneity Model; M4), 총 4가지 모형을 추정하고 값을 비교하였는데, 결과적으로 규모와 선호 이질성을 모두 고려한 일반화된 혼합로짓모형의 경우가 가장 적합한 모형으로 나타났다. 이 때 각각의 확률분포는 삼각분포로 가정하였다. 혼합로짓모형을 기준으로 각각의 모형 별 통행시간과 통행비용의 탄력성을 비교한 결과, 규모이질성 모형은 모든 대안에서 음의 값을 가지고 있으므로 혼합로짓모형보다 통행시간과 통행비용 탄력성이 모두 낮은 것으로 나타났다. 다항로짓모형의 경우에는 모든 대안에서 통행시간 탄력성이 높았으나, 일반화된 혼합로짓모형은 승용차 통행을 제외하고 통행시간 탄력성이 혼합로짓모형보다 높게 나타났다.

Zito and Salvo(2012)는 이탈리아 남부 도시인 팔레르모(Palermo)를 대상으로 대중교통 도착시간에 대한 정보와 정보의 정확성이 이용자들이 수단을 선택할 때 어떤 영향을 미치는 지에 관해 연구하였다. 특정 시나리오에 관한 설문조사를 통해 승용차와 대중교통의 수단선택 로짓모형을 추정하였다. 결과적으로 잠재계층모형의 Log-likelihood와 Rho-squared 값이 더 크게 나와 설명력이 높은 것을 알 수 있었다. 대기시간과 통행시간의 시간가치는 다항로짓모형이나 잠재계층모형이 비슷하지만, 정보의 정확성에 관해서는 시간가치가 큰 것으로 보아 대중교통의 도착 정보가 수단선택 시 중요한 영향을 미치는 것으로 나타났다.

Leon and Miguel(2017)은 아프리카 시에라리온(Sierra Leone)의 수도 프리타운(Freetown)과 룽기공항(Lungi International Airport)을 이동하는 4개의 수단에 관한 수단선택 로짓모형을 혼합로짓모형으로 추정하였다. 통행의 완료여부와 전체 통행비용에 대해서 제한된 삼각분포를 가정하였다. 정규분포나 로그정규분포처럼 극한값을 가져 한쪽으로 치우친 분포를 갖지 않고, 연속적이고 대칭적인 분포를 가지는 제한된 삼각분포로 통행비용의 분포를 가정하였다. 혼합로짓모형과 조건부로짓모형에 의한 생명 통계적 가치(Value of Statistical Life)를 살펴보면, 혼합로짓모형을 사용할 때가 비용 격차가 적게 나타났다. 이를 통해 개인의 생명 통계적 가치금액과 사망위험을 줄이기 위한 정책적 지불의사금액을 혼합로짓모형으로 추정하는 경우가 더 효율적으로 예상하였다.

Tirachini et al.(2017)의 연구에서는 수단선택 시 지하철 내의 혼잡도를 주요한 변수로 고려하여 수단선택 로짓모형을 추정하였다. 통행시간과 입석과 좌석의 혼잡도에 관해 로그정규분포로 가정하여 혼합로짓모형을 추정하였고, 잠재계층모형에서는 여성일수록, 나이가 많을수록, 저소득자의 통행자일수록 혼잡에 민감한 것으로 나타났다. 다항로짓모형과 혼합로짓모형의 경우에는 혼잡도에 대한 민감도가 크게 차이나지 않았으나, 다항로짓모형과 잠재계층모형을 비교해 보았을 때는 개인의 선호다

양성에 의해 혼잡도가 수단선택에 많은 영향을 미치는 것으로 나타났다.

<표 4> 선호다양성을 고려한 수단선택 로짓모형 간 비교에 관한 연구

선행연구	자료 형태	모형	모형 비교방법			우월한 모형
			Rho	시간가치	탄력성	
정진혁 외(2007)	가구통행 실태조사	MNL/MXL	○	○		MXL
이준규 외(2015)	SP	MNL/MXL	○	○		MXL
조신형 외(2017)	SP	MNL/MXL	○	○		MXL
Greene and Hensher(2003)	SP	MNL/MXL/LCM	○	○	○	△
Lee et al.(2003)	SP	MNL/MXL/LCM /LC-MXL	○			LC-MXL
Viton(2004)	RP	MNL/MXL		○		MXL
Shen(2009)	SP	MNL/MXL/LCM	○	○	○	LCM
Greene and Hensher(2010)	SP	MNL/MXL	○	○	○	MXL
Zito and Salvo(2012)	SP	MNL/LCM	○	○		LCM
Leon and Miguel(2017)	SP	CL/MXL		○		MXL
Tirachini et al.(2017)	SP	MNL/MXL/LCM		○		LCM

주: MNL(Multinomial Model), CL(Conditional Logit Model), MXL(Mixed Logit Model), LCM(Latent Class Model), LC-MXL(Latent Class-Mixed Logit Model)
△는 어떠한 모형이 더 우월하다고 할 수 없다고 한 경우를 의미한다.

<표 4>는 선호다양성을 고려한 수단선택 로짓모형을 비교 연구한 선행연구들을 정리한 표이다. 대부분의 연구에서 각각의 모형에 대해 Rho-squared 값으로 모형의 적합성을 판단하였고, 모형 간 비교를 하기 위해 변수의 계수로 시간가치와 탄력성을 계산하였다. 모든 연구가 선호다양성을 고려한 모형이 그렇지 않은 모형보다 우월한 모형이라고 결론지었지만, 혼합로짓모형과 잠재계층모형 모두 장단점을 가지고 있는 모형이라 어떤 모형이 더 우월한 지에 관해서는 연구자가 어떠한 기준으로 모형을 추정하고 해석하느냐에 따라 달라지는 것으로 나타났다.

<표 5> 개인의 선호다양성을 고려한 수단선택 로짓모형 선행연구 고찰

선행연구	저자	자료형태	모형	확률분포	설명변수
하나의 모형	배운경 외(2010)	SP/RP	LCM	-	차내시간, 접근시간, 통행비용, 개인 및 사회경제적 특성
	배운경 외(2011)	가구통행실태조사	LCM	-	통행시간, 통행비용, 개인 및 사회경제적 특성
	추상호(2012)	가구통행실태조사	MXL	정규분포	통행시간, 통행비용 , 개인 및 사회경제적 특성
	Hensher and Greene(2001)	SP/RP	MXL	로그정규분포	통행시간, 통행비용 , 톨게이트비용, 도로 형태
	Atasoy et al.(2011)	SP/RP	LCM	-	통행시간, 통행비용, 개인 및 가구 특성
	Vij et al.(2013)	Travel Diary Surveys	LCM	-	통행시간, 개인 및 가구 특성
둘 이상 의 모형	정진혁 외(2007)	가구통행실태조사	MNL/MXL	균등분포, 정규분포	통행시간, 통행비용 , 개인 및 사회경제적 특성
	이준규 외(2015)	SP	MNL/MXL	삼각분포	통행시간, 운임 , 운항횟수, 안전성, 면세점 이용가능 여부
	조신형 외(2017)	SP	MNL/MXL	정규분포, 삼각분포	통행시간, 통행비용
	Greene and Hensher(2003)	SP	MNL/MXL/LCM	삼각분포	통행시간, 통행비용
	Lee et al.(2003)	SP	MNL/MXL/LCM /LC-MXL	정규분포	통행시간 , 지체시간, 지체시간 정보
	Viton(2004)	RP	MNL/MXL	균등분포	차내시간, 대기시간, 접근시간, 통행비용/세후 임금
	Shen(2009)	SP	MNL/MXL/LCM	정규분포	차내시간, 접근시간, 지체시간 , 빈도, 통행비용, 비환경적영향, 지역 자연환경, 네트워크접근성, 개인 및 사회경제적 특성
	Greene and Hensher(2010)	SP	MNL/MXL	삼각분포	차내시간, 대기시간, 접근시간, 하차시간, 통행비용
	Zito and Salvo(2012)	SP	MNL/LCM	-	대기시간, 통행시간, 정보의 정확성, 개인 및 사회경제적 특성
	Leon and Miguel(2017)	SP	CL/MXL	제한된 삼각분포	통행완료여부, 통행비용 , 안락감, 소음정도, 혼잡도, 응답자의 지위
	Tirachini et al.(2017)	SP	MNL/MXL/LCM	로그정규분포	통행시간, 혼잡도(입석과 좌석) , 개인 및 사회경제적 특성

주: MNL(Multinomial Model), CL(Conditional Logit Model), MXL(Mixed Logit Model), LCM(Latent Class Model), LC-MXL(Latent Class-Mixed Logit Model)

분포를 가정한 설명변수는 Bold처리하였음.

제 3 절 시사점

수단선택 모형에 관련된 선행연구는 국내외적으로 많이 연구되어 있으며, 일반적인 로짓모형의 한계점을 개선한 혼합로짓모형이나 잠재계층모형에 관해서 지속적으로 연구되고 있는 상황이다.

하지만, 국내에서는 혼합로짓모형과 잠재계층모형으로 수단선택 모형을 추정한 연구가 그리 많지 않으며, 특히 일반적인 로짓모형과 선호다양성을 고려한 혼합로짓모형간의 비교연구는 국내 연구를 통해 최근 들어서야 많은 관심을 받고 있는 것을 알 수 있었다. 그러나 국외에서는 로짓모형과 혼합로짓모형 뿐만 아니라 잠재계층모형까지 고려하여 모형을 비교한 연구가 다수 존재한다. 혼합로짓모형의 분포는 주로 정규분포와 삼각분포를 이용하였으며, 각각의 분포의 특징에 따라 가정하는 분포가 달랐다. 이들 연구에서 공통적으로 도출된 결과로는 선호다양성을 고려한 모형이 일반적인 로짓모형보다 더 우월하며, 잠재계층모형이 혼합계층모형보다 더 우월한 결과를 보이고 있다. 하지만, 두 모형간의 우월성이 절대적이지는 않기 때문에 다양한 연구를 통해 확인해보아야 한다는 게 이들 연구의 결론이다.

또한, 수단선택 로짓모형 간 비교연구에 대한 선행연구의 대부분은 SP조사로, 실제 발생한 사건이 아니라 가상의 시나리오를 통해 개인의 선호를 조사하는 것이기 때문에 실제 행동과 오차가 발생할 수 있다.

이러한 점에서 본 연구는 다항로짓모형, 혼합로짓모형, 잠재계층모형 세 가지 모형간의 비교연구를 실제 통근자의 통행을 반영한 가구통행실태조사 자료를 이용한다는 점에서 국외연구와 차별성을 가지고, 가구통행실태조사 자료에서 승용차운전자와 승용차동승자를 구분하여 조사한 것을 이용하여, 승용차운전과 승용차동승을 다른 선택 대안으로 보았다. 또한 아직까지 국내에서 수단선택 로짓모형으로 세 가지 모형을 전부 추정하고 비교하는 연구가 없다는 점에서 의의가 있다.

제 3 장 연구방법론의 정립

제 1 절 수단선택 로짓모형의 설정

가구통행실태조사 자료는 개인의 사회·경제적 특성뿐만 아니라 응답자가 선택한 교통수단과 통행시간, 통행비용 등의 정보를 알 수 있다. 본 연구에서는 수단선택 로짓모형의 가장 중요한 변수로 통행시간과 통행비용만을 설명변수로 사용하고, 통행시간의 경우에는 차내시간과 차외시간을 합친 총 통행시간을 사용한다. 통행비용의 경우에는 2010년을 기준으로 수단별 요금을 다르게 적용한다. 승용차의 경우에는 유류비, 주차비, 통행료, 유지관리비를 모두 고려하고, 버스는 기본요금 900원을, 지하철은 기본요금 900원에 거리에 따른 추가요금을 고려하여 적용한다.

개인의 사회·경제적 특성으로 가구통행실태조사 자료에서 1부터 6까지 범위로 가구의 월평균 소득을 제공하는데, 소득에 따라 동일한 통행비용이더라도 비용에 대한 실질적 체감이 다를 것으로 보고, 통행비용을 통행시간과 같은 단위인 분당 소득으로 나누어 고려한다. 자료가 범위로 제공되기 때문에, 소득계층의 대표 소득으로 중앙값을 이용하고 1과 6의 경우에는 중앙값 대신 각각 80만원과 1,200만원을 대표 소득으로 보았다. 계층별 대표 소득을 고용노동부(2010)에서 제공하는 2010년 월평균근로시간인 182.7시간으로 나누어 소득 계층별 분당 소득을 산정한다. 산정된 계층별 분당 소득은 <표 6>과 같다.

<표 6> 소득계층 별 적용 분당 소득

소득계층	1	2	3	4	5	6
중앙값	80만원	150만원	250만원	400만원	750만원	1,200만원
분당소득	72.98원	136.84원	228.06원	364.90원	684.18원	1,094.70원

수단선택 로짓모형을 추정하기 위해서 가구통행실태조사를 통해 알 수 있는 개인의 사회·경제적 속성에 대한 값들을 더미변수로 추가하였다. 여러 가지의 개인의 사회·경제적 속성이 있지만, 그 중에서 본 연구에서는 성별, 승용차미보유가구원, 고령자와 근무형태에 관한 속성을 더미변수에 포함시켰다.

<표 7> 변수 설명 및 표기 방법

변수	구분	표기
종속변수	교통수단	j
설명변수	통행시간 (분)	TT_j
	승용차	유류비
		주차비
		통행료
	통행비용	유지관리비
	시내버스	버스요금
	도시철도	기본요금
		추가요금
	분당소득	TCT_{nj}
		1(100만원 미만) 72.98원
		2(200만원 미만) 136.84원
		3(300만원 미만) 228.06원
		4(500만원 미만) 364.90원
		5(1,000만원 미만) 684.18원
		6(1,000만원 이상) 1,094.70원
사회·경제적 특성	성별	X_{gender}
	승용차 미보유 가구원	X_{noncar}
	고령자	X_{old}
	근무형태	$X_{employee}$
상수	대안특성상수	ASC_j

<표 8> 개인의 사회·경제적 속성 더미변수 설정

설명변수	설정	기준
TT_j	대안의 통행시간 (All)	
TCI_{nj}	분당 소득 별 통행시간 (All)	
X_{gender}	성별 더미변수 (Bus)	여성 1; 남성 0
X_{noncar}	승용차미보유가구원 더미변수 (Bus, Subway)	20대 1; 그 외 0
X_{old}	고령자 더미변수 (Subway)	65세 이상 1; 그 외 0
$X_{employee}$	근무형태 더미변수 (Subway)	전일제 1; 그 외 0
ASC_j	대안 특성 상수(Driver 제외)	

개인의 사회·경제적 속성에서 여성의 경우 남성보다 직장 근처에 거주하고¹⁰⁾ 혼잡한 지하철에서의 안전에 민감하므로, 접근성이 높고 지하철보다 덜 혼잡한 버스를 더 많이 이용할 것으로 예상된다. 20대의 경우에는 가구주의 비율이 낮고, 승용차를 보유하고 있는 사람이 적기 때문에 승용차를 보유한 가구더라도 승용차가 실제 대안이 되기 어려울 것이므로 20대의 그룹을 구분하였다. 고령자는 서울시 지하철 무임승차혜택을 받으므로, 다른 대안보다 지하철을 더 선호할 것으로 예상되므로, 무료운임 혜택을 받는 65세 이상 고령자그룹을 구분하였다. 또한 본 연구의 대상이 통근자임을 고려하여, 오전 칠두시에 출근하는 전일제 근무자에게는 정시성이 수단선택 시 많은 영향을 미칠 것으로 기대되어 근무형태를 전일제와 그 외로 구분하고, 지하철에만 더미변수로 포함하였다.

기본 다항로짓모형은 식 (10)과 같이 대안특정상수와 설명변수로는 통행시간, 분 단위 소득으로 나눈 통행비용과 개인의 사회·경제적 특성, 종속변수는 대안 j 의 총 효용의 형태로 나타난다. 설명변수의 총 통행시간과 분당 소득으로 나눈 총 통행비용의 경우에는 표준편차의 값이 커서 모델의 설명력을 위해 통계적 보정으로 각각의 설명변수에 로그를 취한 값을 사용한다.

10) 진중현(1994) 연구의 결과

$$U_n(j) = ASC_j + \beta_{TT} \log TT_j + \beta_{TC} \log TCI_{nj} + \beta_k X_{nk} + \varepsilon_j \quad \cdots(10)$$

여기서, $U_n(j)$: 개인 n 의 대안 j 에 대한 총 효용

ASC_j : 대안 j 의 대안특정상수

TT_j : 대안 j 의 총 통행시간

TCI_{nj} : (대안 j 의 총 통행비용)/(개인 n 의 분당 소득)

X_{nk} : 개인 n 의 사회·경제적 특성

β_{TT} , β_{TC} : 통행시간 계수, 통행비용 계수

β_k : 속성변수 k 의 계수

ε_j : 대안 j 의 확률적 효용

개인 n 이 대안 j 를 선택할 확률은 식 (11)과 같으며, 각각의 대안 별로 효용을 비교하여 가장 효용이 높은 수단을 선택한다.

$$P_n(j) = \frac{e^{U_n(j)}}{\sum_{j'=1}^J e^{U_n(j')}} \quad \cdots(11)$$

개인의 선호 다양성을 고려한 혼합로짓모형과 잠재계층모형은 기본 다항로짓모형의 식에 확률변수를 곱하거나 각각의 계층으로 구분한 형태로 모형의 식이 나타난다.

혼합로짓모형의 경우에는 식 (13)에서처럼 β'_k 를 모수가 θ 인 확률밀도함수 $f(\beta|\theta)$ 를 가지는 확률변수이며, 선행연구에서도 볼 수 있듯이 변수의 확률분포를 주로 정규분포(Normal Distribution), 로그정규분포(Lognormal Distribution), 삼각분포(Triangular Distribution), 균등분포(Uniform Distribution)등을 따른다고 가정하고 각각의 모형을 추정하는

것으로 나타났다.

$$U_n'(j) = ASC_j + \beta_{TT}' \log TT_j + \beta_{TC}' \log TCI_{nj} + \beta_k X_{nk} + \varepsilon_j \quad \cdots(12)$$

$$P_n(j) = \int \frac{e^{U_n'(j)}}{\sum_{j'=1}^J e^{U_n'(j')}} f(\beta|\theta) d\beta \quad \cdots(13)$$

일반적으로 가장 많이 사용되는 분포는 정규분포이며, 중심극한정리를 기반으로 한다. 정규분포는 부호에 제약이 없이 대칭적인 분포를 띄고 있는 것이 특징이지만, 최댓값과 최솟값이 정해지지 않아 극한값을 갖고 있으며, 양과 음의 값을 모두 가지고 있어 동일한 변수가 다른 방향으로 개인의 효용에 영향을 미치는 특징이 있다. 정규분포와 달리 자연로그를 취해 모든 통행자에게 동일한 방향으로 효용에 영향을 미치는 로그정규분포는 매우 긴 꼬리를 가지고 있어 다른 분포에 비해 과대 추정되는 문제가 있다. Hensher and Greene(2001)의 연구를 보면, 정규분포, 삼각분포, 균등분포의 통행시간가치는 유사하게 나타났으나, 로그정규분포는 다른 분포보다 통행시간가치가 3배 이상 높게 나타났다. 따라서 통행시간가치를 추정할 때는 로그정규분포가 적합하지 않은 분포임을 알 수 있다. 이러한 문제점을 해결하기 위해 연속적이고 대칭적이며, 최빈값을 기준으로 최솟값과 최댓값의 경계가 있어 극단적인 값을 갖는 것을 방지하는 삼각분포를 선행연구에서 많이 이용하는 것을 알 수 있었다.

혼합로짓모형에서 미지의 확률분포를 연구자가 가정할 수 있지만, 그 가정이 주관적이라는 한계를 고려하여 본 연구에서는 모든 계수의 분포를 가정하는 것이 아니라, 가장 중요한 변수인 통행시간과 통행비용의 계수만 확률분포로 가정한다. 또한, 선행연구에서 이용하는 확률분포 중에서 가장 많이 사용하는 정규분포와 극단적인 값을 갖는 것을 방지하는

삼각분포를 이용하지만, 본 연구에서는 통행시간과 통행비용의 변화가 모든 이용자 효용에 동일한 방향으로 영향을 줄 것으로 예상되기 때문에, 한 방향의 분포만 고려하기 위해 0에서 절단된 정규분포와 절단된 삼각분포의 분포만을 이용한다.

잠재계층모형은 특정 기준에 따라 계층을 구분하고, 각 계층은 서로 비슷한 특징을 가지지만, 계층 간에는 서로 다르다는 특징을 이용하여 모형을 추정한다. 계층 s 의 변수 별 계수 값은 동일하지만, 계층 간 계수 값은 다르기 때문에, 계층 별로 수단선택 로짓모형을 각각 추정하여 계수 값을 비교한다.

$$U_n(j) = ASC_j + \beta_{S_{TT}} \log TT_j + \beta_{S_{TC}} \log TCI_{nj} + \beta_k X_{nk} + \varepsilon_j \quad \cdots(14)$$

$$P_s(j) = \sum_{s=1}^S p_s \left(\frac{e^{U_n(j)}}{\sum_{j'} e^{U_n(j')}} \right) \quad \cdots(15)$$

식 (14)은 계층 s 에 속한 개인 n 의 대안 j 를 선택했을 때의 총 효용이고, 식 (15)는 개인 n 이 계층 s 에 속할 확률 p_s 를 다항로짓모형의 선택확률 식에 곱한 형태이다.

제 2 절 로짓모형 간 비교 분석 방법

1) 모형의 적합성

본 연구에서 추정된 모형의 적합성(Goodness of Fitness)을 판단하는 기준으로 계수의 부호가 직관과 같은 방향을 보이는 지를 확인하여 계수의 적합성을 확인하고, McFadden이 제안한 ρ^2 (Rho-squared) 값을 사용하여 모형이 적합한 지를 확인한다. 마지막으로 추정에 사용되지 않

은 표본을 이용하여 추정된 효용 식이 적절한 지 확인하는 과정을 거쳐 모형의 적합성을 판단하고자 한다.

본 연구에서 가장 중요한 변수는 통행시간과 통행비용으로, 통행시간과 통행비용은 직관적으로 값이 증가하면 개인의 효용은 감소할 것으로 예상된다. 따라서 통행시간과 통행비용의 부호는 음의 부호로 추정되어야 직관에 부합하게 추정된 수단선택 로짓모형이라 할 수 있다.

로짓모형의 경우 종속변수가 선택대안과 비선택대안 두 개로 나뉘는 이산형 변수이므로 회귀모형에서 사용하는 R^2 값을 이용하는 경우에 편차가 발생하므로,¹¹⁾ 본 연구에서는 McFadden이 제안한 ρ^2 값을 이용하며, 이는 선형 회귀모델에서 모형의 적합도를 판단하는 R^2 값과 동일한 기능을 하는 것으로 볼 수 있다. ρ^2 는 식 (16)에서와 같이 상수항만 있는 모형(Intercept-only Model) L_{null} 과 전체 모형(Full Model) L 을 Log-likelihood Kernels를 이용하여 산출한다.¹²⁾

$$\rho^2 = 1 - \frac{\ln L}{\ln L_{null}} \quad \dots(16)$$

ρ^2 도 회귀모형의 R^2 과 마찬가지로 0에서 1사이의 값을 가지며, 1에 가까운 값일수록 적합한 모형을 나타내지만, 일반적으로 0.2에서 0.4의 값만 가지더라도 좋은 적합도를 가지는 것이다.

추정된 수단선택 로짓모형이 제대로 추정된 지 확인하기 위해, 실제 이용한 수단과 최대효용의 수단의 일치 여부를 확인하여 실제로 선택한 수단의 50% 이상이 최대효용수단이라면 적합한 모형이라고 판단한다.

각각의 수단을 승용승합, 승용차동승, 버스, 지하철에 대해 1부터 4까지 각각 번호를 매김하고, 효용이 가장 높은 수단과 실제 이용한 수단

11) 최영은(2017)을 참고하였다.

12) 이희연 외(2012)의 로지스틱 회귀모형에서 개념을 참고하였다.

이 동일한 지를 확인한다.

$$CM = \frac{U_{\max, mn}}{C_{mn}} \quad \dots(17)$$

여기서, CM : 가장 효용이 높은 수단을 이용한 지 일치여부

$U_{\max, mn}$: 여러 수단 중 가장 큰 효용의 수단 번호

C_{mn} : 실제 이용한 수단의 번호

실제 이용한 수단이 가장 큰 효용의 수단인 경우 $CM=1$, 실제 이용 수단이 가장 큰 효용의 수단이 아닌 경우는 $CM \neq 1$ 의 값을 가진다.

$$VC = \left(\sum_{n=1}^N CM \right) / N \times 100 \quad \dots(18)$$

가장 효용이 높은 수단을 이용한 지를 확인하고, 모든 이용자에 대해 그 값을 다 더해 전체 표본 수로 나눈 VC 값이 50%이상이 나오면, 추정된 수단선택 모형은 적합하다고 설정하였다.

또한, 적용 표본의 실제 수단분담률과 추정된 식에 의해 산정된 추정분담률을 비교하여 추정분담률이 실제분담률과 가장 비슷한 모형이 가장 설명력이 높은 모형이라 보았다.

AUR (Actual Usage Rates)는 적용 표본의 이용자의 실제 수단분담률을 의미하고, EUR (Estimation Usage Rates)는 그 표본으로 추정된 모형 식에서 최대효용수단을 확인한 추정분담률을 의미한다.

$$AUR = (\text{수단별 이용객 수} / N) \times 100 \quad \dots(19)$$

$$EUR = (\text{수단별 이용객 수} / N) \times 100 \quad \dots(20)$$

2) 모형의 추정결과

모형의 추정결과를 비교하기 위해서 선행연구와 마찬가지로 모형별 계수추정치를 이용하여 통행시간가치(Value of Time, VOT)와 통행시간, 통행비용의 수요탄력성을 구한다.

통행시간가치는 이용자가 어떤 수단을 선택하느냐에 따라 이용자 개인이 느끼는 시간 가치가 다르다는 한계대체율법(Marginal Rate of Substitution)으로 통행시간을 산정한다.¹³⁾ 한계대체율법은 개별행태모형의 통행시간과 통행비용의 계수를 통해 통행시간을 산정하는데, 본 연구에서 다루는 세 가지 모형은 모형의 종류에 따라 통행시간과 통행비용의 계수 값이 하나의 값 또는 확률분포로 나타나기 때문에 통행시간가치를 비교함으로써 모형 간 추정결과 비교가 가능하다.

일반적인 통행시간가치는 식 (21)로 추정이 가능하며, 일반적으로 통행시간과 통행비용의 계수 비로 통행시간가치를 구하지만, 본 연구에서는 가구 소득으로 나눈 통행비용을 사용하므로 평균 가구소득을 곱하고, 통행시간 단위를 맞추기 위해 60을 곱한 값이 통행시간가치를 나타낸다.

$$VOT_j = \frac{\partial U_j / \partial \log TT_j}{\partial U_j / \partial \log (TCI_{nj})} = \frac{\beta_{TT}}{\beta_{TC}} \times INC_{avg} \times 60 \quad \dots(21)$$

여기서, VOT_j : 대안 j 에 대한 통행시간가치

U_j : 선택 대안에 대한 총 효용

TT_j : 선택 대안 j 에 대한 통행시간

TCI_{nj} : (대안 j 의 총 통행비용)/(개인 n 의 분당 소득)

13) 정진혁 외(2007) 연구를 참고하였다.

β_{TT} β_{TC} : 통행시간 계수, 통행비용 계수

INC_{avg} : 평균 가구소득

혼합로짓모형의 경우에는 추정된 확률변수의 평균과 표준편차로 모집단의 분포를 먼저 추정하고, 이를 이용하여 통행시간가치를 추정한다. 추정된 계수의 평균과 표준편차를 이용하여 통행시간과 통행비용의 효용을 구할 수 있고, 각각의 효용함수 식은 식 (22)와 (23)이다. 식 (24)는 모집단의 평균 통행시간가치를 구하는 식이고, 여기에서 N 은 전체 모집단 표본의 수, c 는 분당 소득 대비 통행비용으로, 본 연구에서 말하는 모든 통행비용은 분당 소득으로 나눈 통행비용을 의미하기 때문이다.

$$U_{time} = \beta_{time} + \sigma_{time} \times t \quad \dots(22)$$

$$U_{cost} = \beta_{cost} + \sigma_{cost} \times c \quad \dots(23)$$

$$VOT'_j = \left(\sum_{n=1}^N \frac{\beta_{time} + \sigma_{time} \times t}{\beta_{cost} + \sigma_{cost} \times c} \right) / N \times INC_{avg} \times 60 \quad \dots(24)$$

잠재계층모형의 경우에는 계층별로 평균소득의 값이 다르게 나타나므로, 계층별로 통행시간가치를 각각 구해야만 한다.

$$VOT_{s,j} = \frac{\partial U_j / \partial \log TT_j}{\partial U_j / \partial \log (TCI_{nj})} = \frac{\beta_{TT}}{\beta_{TC}} \times INC_{s,avg} \times 60 \quad \dots(25)$$

모형의 추정결과를 비교하는 다른 방법으로 통행시간과 통행비용에 대한 수요탄력성을 살펴본다. 일반적으로 경제학에서 수요탄력성이라 함은 가격이 변화함에 따라 수요가 변하는 정도를 보는 것이다. 그런 의미로 통행시간탄력성은 통행시간이 한 단위 증가하였을 때 선택 대안의 수요가 변화하는 지를 보는 것이고, 통행비용탄력성은 통행비용이 한 단위

증가하였을 때 선택 대안의 수요의 변화를 보는 것이다.

식 (26)은 통행시간탄력성으로 T_j 와 ΔT_j 는 각각 특정수단의 시간과 그 변화량을 말하며 Q_j 와 ΔQ_j 는 수요량과 시간의 변화량에 따른 수요의 변화량을 의미한다.

$$E_{TT} = - \frac{\Delta Q_j}{\Delta T_j} \cdot \frac{T_j}{Q_j} \quad \dots(26)$$

마찬가지로 식 (27)은 통행비용탄력성으로 C_j 와 ΔC_j 는 특정수단의 비용과 비용의 변화량을 말하고 Q_j 와 ΔQ_j 는 수요량과 비용의 변화량에 따른 수요의 변화량을 의미한다.

$$E_{TC} = - \frac{\Delta Q_j}{\Delta C_j} \cdot \frac{C_j}{Q_j} \quad \dots(27)$$

제 4 장 자료의 구축

제 1 절 서울시의 현황 분석

1) 서울시

본 연구에서는 통근자의 수단선택 모형을 추정하기 위해 가구통행실태조사 자료를 이용하였다. 정부에서 5년마다 시행하는 가구통행실태조사는 전국을 대상으로 하며, 실제 개인의 통행을 기반으로 설문조사가 진행된다. 연구에 사용되는 자료는 가장 최신 자료인 2010년 가구통행실태조사 자료를 이용한다.¹⁴⁾ 2010년 가구통행실태조사 자료의 전체 표본 수는 1,048,574개이며, 그 중 서울시 내 통행은 541,230개로 집계되었다.

서울시 내 통행목적별 통행량과 통행수단별 통행량은 <표 9>와 <표 10>에 나타나있다. 통행 목적 중에서는 귀가를 제외하고는 출근 통행이 19.17%로 가장 높았고, 그 다음이 등교로 11.70%를 차지하고 있었다. 통행 수단별로는 도보를 제외하고는 지하철 18.98%, 버스 18.58%, 승용승합 15.96%, 승용차동승 5.79%, 기타 4.62%, 택시 1.14% 순으로 비율을 차지하는 것으로 나타났다.

<표 9> 서울시 통행목적별 통행량

(단위 : 통행, %)

목적	출근	등교	업무	쇼핑	학원	여가	기타	귀가	합계
통행량	103,738	63,313	14,555	16,800	31,572	21,580	45,826	243,846	541,230
비율	19.17	11.70	2.69	3.10	5.83	3.99	8.47	45.05	100.00

14) 2015년 가구통행실태조사는 현재 전수화 중에 있어(2017년 10월 기준) 자료가 배포되지 않았기 때문에 가장 최신 자료인 2010년 자료를 사용하였다.

<표 10> 서울시 통행수단별 통행량

(단위 : 통행, %)

수단	승용승합	승용차동승	버스	지하철	택시	도보	기타	합계
통행량	86,385	31,325	100,551	102,737	6,146	189,062	25,024	541,230
비율	15.96	5.79	18.58	18.98	1.14	34.93	4.62	100.00

주: 버스 통행량에 고속버스 이용량 100통행 포함.

지하철 통행량에 고속철도(KTX) 이용량 46통행 포함.

기타 : 소형화물차, 중대형화물차, 오토바이, 자전거, 기타통행 포함.

가구통행실태조사 자료에 의하면, 서울시 전체 가구 중 자동차 미보유가구와 자동차보유가구의 비율은 대략 30%와 70%로 대부분의 가구에서 승용차를 보유하고 있는 것으로 나타났다. 하지만, 대부분이 한 가구당 한 대의 자동차를 보유하고 있으므로, 가족 중 한 명만 승용차를 이용하거나 동승을 할 것이라고 예상된다.

<표 11> 가구당 자동차 보유대수

(단위 : 가구, %)

구분	없음	1대	2대	3대	4대 이상	합계
가구 수	28,982	57,722	10,432	749	177	98,062
비율	29.55	58.86	10.64	0.76	0.18	100.00

통계청에서 제공하는 서울시 가구주의 연령 자료를 살펴보면, 40대가 가장 많고, 50대, 30대 순으로 많다. 그에 반해 20대, 60대, 70대 이상은 대부분이 가구주가 아닌 것으로 나타나 승용차보유가구의 가구원이지만, 이 연령대는 본인은 차량을 소유하고 있지 않을 가능성이 높다.

<표 12> 서울시 가구주 연령

(단위 : 명, %)

연령	20대	30대	40대	50대	60대	70대 이상	합계
수	374,186	769,208	871,521	779,418	481,015	283,634	3,558,982
비율	10.51	21.61	24.49	21.90	13.52	7.97	100.00

자료 : 통계청 「장래가구추계」

2) 서울시 통근자

본 연구에서는 서울시 내 통행자 중 통근목적의 통행자만을 대상으로 분석을 진행하였다. 이 때, 대중교통의 환승에 관해 정확하게 파악하기 어려우므로 환승의 경우는 제외하였으며, 출발지와 도착지가 서울시인 자료만을 서울시 내부 통행으로 보았다.

전체 541,230개의 서울시 내부 통행 중에서 통근통행은 103,738개로 통근통행의 수단분담률은 <표 13>과 같다. 서울시 내 통근통행자들 중에서는 지하철을 이용하는 사람이 29.37%로 가장 많았고, 자가용 이용자가 25.48%로 그 다음으로 많았으며, 특이한 점으로 택시보다 승용차 동승 통행량이 더 많은 것으로 나타났다.

<표 13> 서울시 내 통근통행 수단분담률

(단위 : 통행, %)

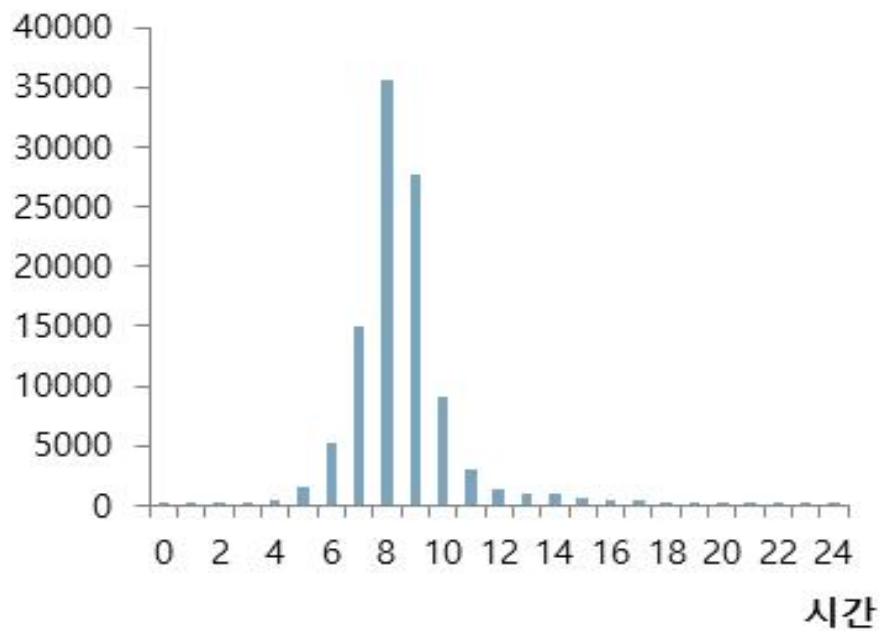
수단	승용승합	승용차동승	버스	지하철	택시	도보	기타	합계
통행량	26,430	3,111	20,530	30,472	986	16,643	5,566	103,738
비율	25.48	3.00	19.79	29.37	0.95	16.04	5.37	100.00

주: 버스 통행량에 고속버스 이용량 23통행 포함.

철도/지하철 통행량에 고속철도(KTX) 이용량 15통행 포함.

기타 : 소형화물차, 중대형화물차, 오토바이, 자전거, 기타통행 포함.

서울시 내 103,738개의 통근통행의 출발시간 분포를 <그림 2>을 통해 살펴보면 오전 7시에서 오전 9시까지가 가장 빈도가 높은 것으로 나타났다. 오전 7시에서 오전 9시까지를 오전 첨두시라 칭하고, 본 연구에서는 가장 통행빈도가 높은 오전 첨두시의 통근통행 자료만을 이용하였다. 오전 첨두시의 서울시 내 통근통행량은 47,598개로 전체 서울시 통근통행의 45.9%가 오전 첨두시에 발생하는 것으로 나타났다.



<그림 2> 통근통행 출발시간 분포

통계청의 자료와 비교하기 위해, 2010 가구통행실태조사 자료의 가구원 자료에서 서울시에 직장이 있는 가구주의 연령을 살펴보면 <표 14>와 같다. 전수자료가 아니라 통계청에서 제공하는 자료와 비율의 차이는 존재하지만 20대, 60대, 70대 이상에서 가구주 비율이 낮게 나타나 이들은 통근 시 수단을 선택할 때 승용차를 본인이 이용할 수 있는 대안이 아님을 알 수 있다.

<표 14> 서울시 통근 가구주 연령

(단위 : 명, %)

연령	20대	30대	40대	50대	60대	70대 이상	합계
수	1,987	15,657	31,080	25,561	10,996	2,246	87,527
비율	2.27	17.89	35.51	29.20	12.56	2.57	100.00

자료 : 2010 가구통행실태조사 자료

제 2 절 자료의 구축 및 특성 분석

1) 자료의 구축

(1) 종속변수

원시자료에서는 총 18개의 통행수단이 조사되었으나, 본 연구에서는 서울시 내 통근통행 수단별 통행량이 가장 높은 세 가지 수단(승용승합, 시내버스, 지하철)과 승용차로 이용할 수 있는 대안인 승용차동승까지 총 4가지 수단을 통근자가 선택할 수 있는 대안으로 선정하였다.

가구현황조사에서 가구별 승용차보유여부를 파악하여 승용차보유가구와 승용차미보유가구 구분하였다. 그 결과, 오전 칠두시 통근 통행을 하는 47,598개의 가구 중, 승용차보유가구는 88.96%, 승용차미보유가구는 11.04%를 차지하는 것으로 나타났다.

<표 15> 승용차 보유/미보유 가구별 통근 통행량

(단위 : 통행, %)

통근통행		승용차보유가구	승용차미보유가구	합계
서울시	통행량	42,344	5,254	47,598
	비율	88.96	11.04	100.00

본 연구에서는 3개 이상의 선택대안을 고려할 것이므로 승용차가 선택 대안이 될 수 있는 승용차보유가구만을 대상으로 선정한다. 승용차보유가구의 선택대안은 승용승합, 승용차동승, 시내버스, 지하철 4가지이며, 승용차동승의 경우에는 통행량이 전체의 3%밖에 차지하지 않지만, 카풀이 과거에 비해 점차 보편화되는 점을 고려하여 승용차동승을 고려하였다. 또한 통근통행 소요시간이 3시간 이상인 자료는 outlier로 보고 제외하였다.

오전 첨두시에 통근통행을 하는 승용차보유가구에 대한 42,344개의 자료 중에서 약 4.2%인 1,780개의 자료를 선택 대안별 수단분담률을 고려하여 랜덤 추출하였다. 표본의 승용차보유가구는 1,780개이며, 그 중에서 승용승합과 철도/지하철의 수단분담률이 거의 40%로 가장 높게 나타났다.

하지만 랜덤 추출한 표본에는 출발지와 도착지의 행정동이 인접하여 대중교통을 이용하는 것이 비효율적인 경우와 2010년 기준 지하철이나 버스 노선이 없어 대중교통으로 접근이 불가능한 지역¹⁵⁾처럼 선택 대안이 없는 지역의 경우도 포함되어 있다. 이러한 경우도 예외적인 경우라고 보고 표본에서 제외하였으며, 그 결과 전체 1,447개의 표본이 구축되었다.

<표 16> 승용차 보유 가구별 통근통행 수단분담률

(단위 : 통행, %)

대안		승용승합	승용차동승	버스	지하철	합계
서울시	통행량	17,323	1,756	6,476	16,789	42,344
	비율	40.91	4.15	15.29	39.65	100.00
랜덤 추출한 표본	통행량	728	74	272	706	1,780
	비율	40.90	4.16	15.28	39.66	100.00
이상치 제외한 표본	통행량	515	50	183	699	1,447
	비율	35.59	3.45	12.65	48.31	100.00

<표 16>에서처럼 구축된 표본의 수단분담률을 살펴보면 지하철이 48.31%로 가장 높고, 승용승합은 35.59%로 나타난다. 이는 이상치의 기준이 대중교통의 접근이 비효율적인 지역이기 때문에 선택대안에 대중교

15) 경춘선(2010년 12월), 신분당선(2011년 10월), 분당선(왕십리~선릉 구간, 2012년 10월), 경의중앙선(공덕~용산 구간 직결 운행, 2014년 12월), 9호선(신논현~종합운동장 구간, 2015년 3월), 우이신설선(2017년 9월) 개통

통이 없어 승용차를 이용한 자료가 제외되어 이러한 결과가 나타난 것으로 보인다. 또한, 이상치를 제외하기 전에 승용승합과 지하철의 분담률이 비슷했던 것으로 미루어보아 이상치의 범위에 2010년 기준 대중교통으로 접근이 불가능한 지역에서는 주로 승용차를 이용해서 이동을 했을 것으로 예상된다.

이상치를 제외한 표본에서 추정과 적용에 사용되는 표본을 수단분담률을 고려하여 각각 70%와 30%로 구분하였으며, 1,013개의 추정표본과 434개의 적용표본으로 구분되었다.

<표 17> 추정과 검증에 사용되는 표본의 수단분담률

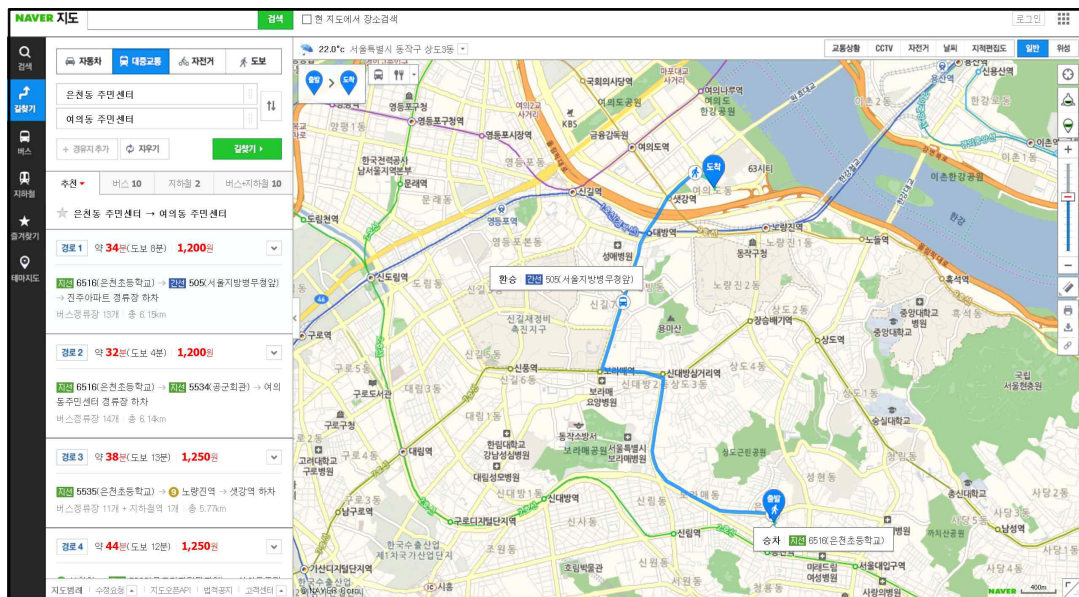
(단위 : 통행, %)

		승용승합	승용차동승	버스	지하철	합계
표본	통행량	515	50	183	699	1,447
	비율	35.59	3.45	12.65	48.31	100.00
추정(70%)	통행량	361	35	128	489	1,013
	비율	35.63	3.46	12.64	48.27	100.00
적용(30%)	통행량	154	15	55	210	434
	비율	35.48	3.46	12.67	48.39	100.00

(2) 통행시간, 통행비용

본 연구에서 가장 중요한 변수는 통행시간과 통행비용이다. 가구통행실태조사에 응답자가 표기한 출발시간과 도착시간의 차이로 통행시간을 구할 수 있다. 하지만, 선택하지 않은 대안의 통행시간과 통행비용은 제공되지 않아 본 연구에서는 네이버 지도 경로검색을 활용하였다. 가구통행실태조사 자료에 제시된 출발지와 도착지의 행정동 주민센터를 각각의 행정동의 기준으로 설정하고, 네이버 지도에서 제시하는 버스와 지하철의 통행시간과 통행비용을 이용하였다.

통근이라는 통행의 특성을 고려하여 통근자는 최소통행시간이 걸리는 경로를 선택할 것이므로, 최소통행시간과 그 때의 통행비용을 사용하였다. 또한 집계되지 않은 승용차의 통행시간은 실시간 도로상황을 반영한 시간을 이용해야하므로, 네이버 지도에서 현재 도로상황을 반영한 통행시간을 제공하는 것을 이용하여 평일 오전첨두시(오전 7시-9시)에 네이버 지도 검색을 실시하여 그 때의 통행시간을 사용하였다.



<그림 3> 네이버 지도 검색 예시

2010년 기준 bus와 지하철의 기본요금은 900원으로 고려하여 현재 금액에서 300원을 차감하였으며, 지하철의 경우에는 거리비례제로 인해 기본 10km를 넘는 이동의 경우에는 초과 5km당 100원씩 추가요금이 부과된 것을 감안하여 그만큼 추가요금을 감안하였다.

승용차의 통행비용은 통행시간 산정 시 제공되는 통행거리에 승용차 운영비용을 곱한 값을 이용하였다. 승용차의 운영비용은 <표 18>과 같이 국가교통DB센터(2006)에서 제공하는 속도별 승용차 운영비용을 이용하였다. 이 자료는 2006년 기준이기 때문에, 유류비 증가율과 물가 상승률을 고려하여 2010년 기준으로 보정한 다음 표본의 평균통근통행속도를 산정하여 운영비용을 적용하였다.

<표 18> 속도별 승용차 운영비용

(단위 : 원/km)

속도(km/h)	유류비 ¹⁾	엔진오일비 ²⁾	타이어비 ²⁾	유지정비비 ²⁾	합계
10	199.51	7.55	0.56	9.66	217.28
20	120.30	6.29	1.04	11.42	139.05
30	102.33	5.45	1.60	13.52	122.90
40	92.42	4.61	2.33	14.05	113.41
50	94.39	4.61	2.96	15.80	117.76
60	98.00	4.61	3.77	16.68	123.06
70	103.39	4.61	4.64	17.56	130.20
80	105.96	4.19	5.61	19.31	135.07
90	115.26	3.77	6.79	19.84	145.66
100	124.23	4.61	8.07	21.08	157.99
110	134.42	5.87	9.84	19.84	169.97
120	150.71	8.80	11.68	25.46	196.65

자료 : 2006 광역권 여객통행조사

주: 1) 유류비 증가율 16.71%를 보정하여 사용함.

2) 물가 상승률 14.17%를 보정하여 사용함.

표본의 평균통근통행속도는 표본을 구축에 관한 절에서 살펴보도록 하겠다. 수도권교통본부(2012)에 의하면 2010년 서울시 평균통근통행시간은 38.0분이고, 평균통근통행거리는 11.10km로 나타나며 따라서 평균통근통행속도는 17.62km/h이다. 표본과 비교해보았을 때, 표본의 평균통행속도는 17.03km/h로 수도권교통본부(2012)의 자료와 거의 유사하다.

<표 19> 서울시 평균 출근 지표

	평균통행시간	평균통행거리	평균통행속도
서울시 기준	38.0분(0.63시간)	11.10km	17.62km/h
표본	43.4분(0.72시간)	12.26km	17.03km/h

자료 : 여객 기종점통행량(O/D) 전수화 및 장래수요예측 공동조사, 2012

앞에서 언급한 수도권교통본부(2006)의 속도별 승용차 운영비용에서 서울시 평균통행속도의 범위인 20km/h일 때의 139.05원을 사용하며, 계산의 편의를 위해 승용차동승의 경우에는 승용차 운영비용의 1/2값을 이용한다.¹⁶⁾

16) 2010년 가구통행실태조사에 의하면 서울시 통근통행의 평균재차인원은 1.3인/대로 나타났다.(출처: <https://www.ktdb.go.kr/www/selectBbsNttView.do?key=42&bb sNo=7&nttNo=474>)

2) 자료의 특성 분석

가구통행실태조사 자료에서 추출한 표본 중 추정과 적용에 사용되는 표본은 각각 1,013개와 434개이며, 이 자료의 사회·경제적 특성을 살펴보면 <표 20>과 같다. 추정표본과 적용표본의 수단별 사회·경제적 특성과 사회·경제적 특성별 수단분담률에 관한 자세한 수치와 비율은 <부록 1>과 <부록 2>에 수록하였다.

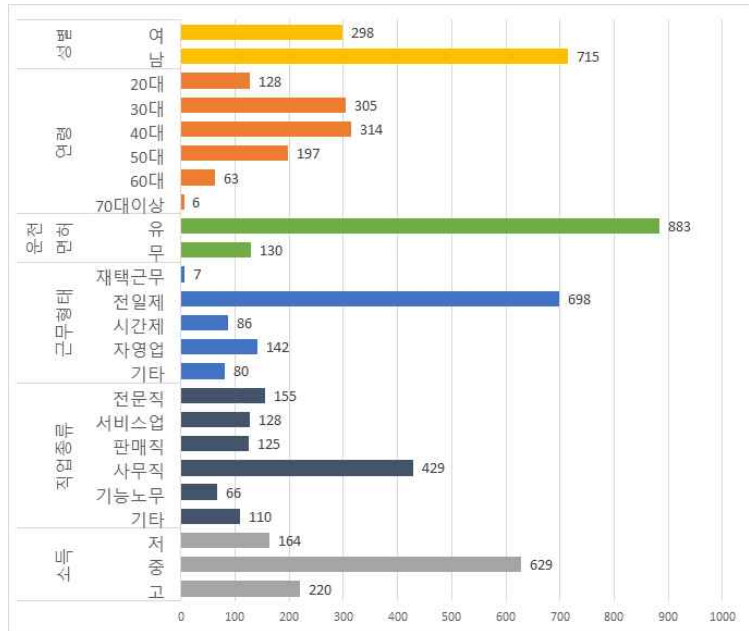
분석에 사용하는 추정 표본의 사회·경제적 특성을 보면 남성이 715명, 여성이 298명으로 남성이 여성보다 많고, 연령에서는 30대와 40대 모두 전체 표본의 30%의 비율로 많이 차지하고 그 다음으로 50대가 많이 차지하는 것을 알 수 있다. 통근자를 대상으로 하기 때문에 30대, 40대와 50대가 주로 경제활동을 하는 연령임을 확인할 수 있다. 운전면허의 유무는 승용차 운전자의 경우 필수요소이므로 전체 표본에서 운전면허를 소지한 표본이 87%로 운전면허 미보유자보다 많다. 근무형태는 전일제가 69%로 나머지 근무형태보다도 월등히 많은 것으로 나타났고, 다른 사람들보다 전일제로 근무하는 사람들이 오전 칠두시에 많이 출근하는 것으로 보인다. 직업종류는 사무직이 가장 많았으며, 그 다음 전문직, 서비스업, 판매직 등이 많았다. 소득은 저소득계층을 월평균 200만원 미만, 중소득을 200만원이상에서 500만원 미만, 고소득계층을 월평균 가구소득이 500만원이상인 계층으로 구분하였으며, 중소득계층이 가장 많은 것으로 나타났다.

적용에 사용할 표본은 추정에 사용하는 표본의 수단분담률을 고려하여 랜덤 추출하였다. 하지만 적용 표본을 구축할 때, 개인의 사회·경제적 특성까지 비슷한 비율로 추출된다면 추정결과를 적용하고 확인하는 데 신뢰성이 떨어진다고 판단하여 적용에 사용하는 표본을 랜덤 추출할 때 추정 표본의 사회·경제적 특성의 비율은 고려하지 않았다. 성별이나 운전면허는 추정 표본과 거의 비율이 비슷하지만, 연령, 근무형태, 직업종류, 소득의 경우 추정 표본과 다른 비율로 랜덤 추출 되었으므로 추정

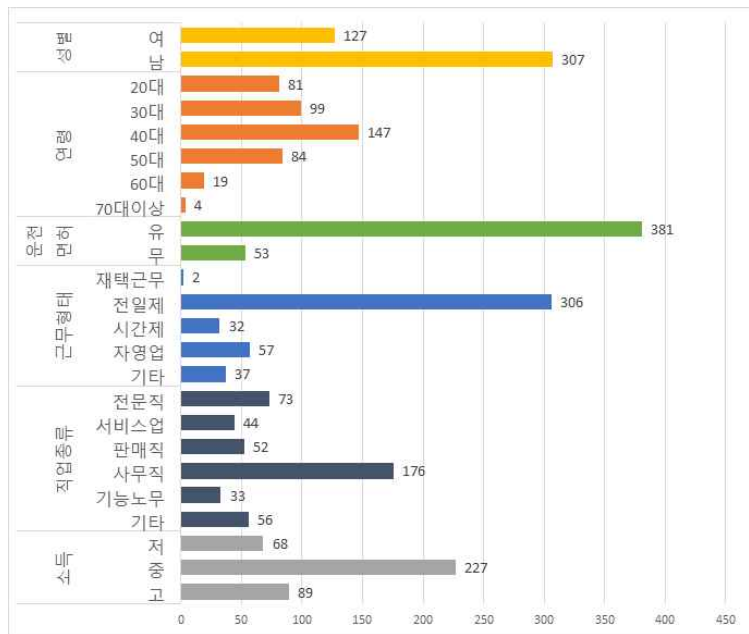
식을 적용하여 추정의 적합성을 확인하는 결과의 의미가 있을 것이다.

<표 20> 표본의 사회·경제적 특성

구분	특성	추정		적용	
		표본 수(명)	비율(%)	표본 수(명)	비율(%)
		1,013	70.01	434	29.99
성별	여성	298	29.42	127	29.26
	남성	715	70.58	307	70.74
연령	20대	128	12.64	81	18.66
	30대	305	30.10	99	22.81
	40대	314	31.00	147	33.87
	50대	197	19.45	84	19.35
	60대	63	6.22	19	4.38
	70대 이상	6	0.59	4	0.92
운전면허	유	883	87.17	381	87.79
	무	130	12.83	53	12.21
근무형태	재택근무	7	0.69	2	0.46
	전일제	698	68.90	306	70.51
	시간제	86	8.49	32	7.37
	자영업	142	14.02	57	13.13
	기타	80	7.90	37	8.53
직업종류	전문직	155	15.30	73	16.82
	서비스업	128	12.64	44	10.14
	판매직	125	12.34	52	11.98
	사무직	429	42.35	176	40.55
	기능노무	66	6.52	33	7.60
	기타	110	10.86	56	12.90
소득	저	164	16.19	68	15.67
	중	629	62.09	277	63.82
	고	220	21.72	89	20.51



<그림 4> 표본의 사회·경제적 특성(추정)



<그림 5> 표본의 사회·경제적 특성(적용)

표본의 수단별 사회·경제적 특성과 사회·경제적 특성별 수단분담률에 관한 자세한 수치와 비율은 <부록 1>에서 확인할 수 있으며, 특이한 점으로는 남성은 승용승합과 지하철을 많이 이용하는 반면, 여성의 경우에는 지하철과 버스의 이용률이 높게 나타났다. 특히 버스와 지하철의 이용률이 22.48%로, 승용승합보다 버스를 더 많이 이용하는 것을 알 수 있었다.

또한, 20대는 사회초년생으로 승용차 보유가구에 속해있지만, 자차를 소유할 가능성이 높지 않아 다른 연령에 비해 대중교통을 많이 이용하는 것으로 보인다. 서울시의 65세 이상 고령자 지하철 무임승차 제도로 인해 60대 이상의 고령자 또한 지하철을 많이 이용하는 것으로 나타났다. 특히 70대 이상의 경우에는 버스의 수단분담률이 0%로, 대중교통을 이용한다면 버스보다는 무료 이용이 가능한 지하철을 선택하는 것을 확인할 수 있다. 운전면허를 보유하지 않은 사람은 버스나 승용차동승보다는 지하철을 더 많이 이용하는 것으로 나타났고, 운전면허를 보유한 사람도 지하철의 이용률이 가장 높은 것으로 나타났다. 또한 저소득계층은 승용차보다는 버스와 지하철을 이용하고, 중소득계층이나 고소득계층 또한 의외로 승용차보다 지하철을 더 많이 이용하는 것을 알 수 있다. 이러한 현상은 통근이라는 목적통행의 특성 상 정시성이 수단선택 시 매우 중요한 요소이기 때문에 운전면허가 있더라도, 고소득계층이더라도 지하철을 이용하는 것으로 보인다.

본 연구에서 가장 중요하게 사용되는 변수인 통행시간과 통행비용에 대한 특성을 살펴보기 위해서 커널밀도(Kernel Density)¹⁷⁾를 이용하여 통행시간과 통행비용의 분포를 확인할 수 있다. 변수의 분포는 <부록 3>에 수록하였다.

17) 커널밀도(Kernel Density)는 히스토그램의 불연속성을 완화시키기 위한 방법이다.

<표 21> 표본의 통행시간과 통행비용 분포

	전체		
	TT	TC	TCI
Minimum	5	102	0
1st Quartile	30	900	2
Midian	40	1,000	3
Mean	43.38	1,185.96	4.41
3rd Quartile	60	1,225	6
Maximum	130	4,083	34
Mode	30	900	2
Std.	17.97	541.68	3.51

주: TT(Travel Time), TC(Travel Cost), TCI(Travel Cost by Income per minute)

통행시간과 통행비용의 분포를 살펴보면, 두 분포 모두 평균값이 중앙값보다 크므로 왼쪽으로 약간 치우쳐 있는 분포임을 알 수 있다. 또한 최빈값은 통행시간은 30분, 통행비용은 900원으로 나타났다.

수단별 통행시간과 통행비용 별 Quartile값과 평균, 표준편차 값은 <표 22>와 같다. 승용승합과 지하철은 평균 통행시간 값이 38.95분과 48.28분으로 중앙값인 40분과 50분보다 조금 작게 나타났고, 이는 통행시간 분포가 약간 오른쪽으로 치우친 것으로 보인다. 그에 반해 승용차동승과 버스의 통행시간은 평균값이 중앙값보다 더 크게 나타났지만, 거의 비슷한 것을 알 수 있다.

<표 22> 수단별 통행시간과 통행비용 별 Quartile

(단위 : 분, 원)

	승용승합		승용차동승		버스		지하철	
	TT	TC	TT	TC	TT	TC	TT	TC
Minimum	5	204	5	102	10	900	10	900
1st Quartile	30	1,225	20	408	30	900	35	900
Midian	40	1,633	30	613	40	900	50	1,000
Mean	38.95	1,591	30.51	623	40.67	900	48.28	1,002
3rd Quartile	50	2,042	40	817	50	900	60	1,100
Maximum	100	4,083	70	1,429	100	900	130	1,500
Std.	17.76	725.20	15.75	321.62	15.97	0.00	17.28	114.02

주: TT(Travel Time), TC(Travel Cost)

통행비용의 경우, 버스나 지하철의 통행비용은 버스의 경우에는 거리에 상관없이 모두 동일한 비용이고, 지하철도 5km당 100원이 추가로 부과되지만 금액이 크지 않아 큰 차이가 없는 것으로 보인다. 하지만, 승용차의 경우에는 이동거리에 따라 통행 비용이 부과되므로 단거리와 장거리 간의 통행비용 표준편차가 크게 나타난다. 승용승합의 경우 통행비용의 평균값이 중앙값보다 작아 통행비용이 오른쪽으로 조금 치우친 분포임을 알 수 있지만, 승용차동승의 경우에는 평균값이 중앙값보다 큰 것으로 나타났다. 이를 통해, 승용차 동승은 장거리보다는 단거리에서 많이 이용하는 수단임을 확인할 수 있고, 동승의 경우 운전자와 목적지가 다르기 때문에 장거리보다는 단거리에서 효율적인 것으로 보인다.

본 연구의 통행시간과 분당 소득 별 통행비용 변수는 통계적 보정을 위해 로그를 취한 값을 사용했고, 로그를 취하기 전보다 표준편차의 값이 감소한 것을 확인할 수 있다. <부록 3>을 참고하면, 로그를 취하기 전에는 한쪽으로 치우친 분포를 보였다면, 로그를 취하고 나니 분포가 거의 정규분포의 형태로 나타났다.

<표 23> 통행시간과 통행비용 변수의 분포

	전체		
	LogTT	LogTC	LogTCI
Minimum	0.6990	2.0090	-0.4283
1st Quartile	1.4771	2.9542	0.3779
Midian	1.6021	3.0000	0.6001
Mean	1.5949	3.0378	0.5150
3rd Quartile	1.7782	3.0881	0.6833
Maximum	2.1139	3.6110	1.5540
Mode	1.4771	0.1746	0.2986
Std.	0.2038	0.1747	0.2989

주: 모든 값은 소수점 다섯째자리에서 반올림하였음.

TT(Travel Time), TC(Travel Cost), TCI(Travel Cost by Income per minute)

제 5 장 모형의 추정 결과

제 1 절 수단선택 로짓모형

1) 다항로짓모형

서울시 통근자의 선호다양성을 고려한 수단선택 로짓모형을 추정하기 위해 본 연구에서는 LIMDEP 9.0, NLOGIT 4.0 프로그램을 이용하여 모형을 추정하였다.

선호다양성을 고려한 수단선택 로짓모형과 선호다양성을 고려하지 않은 모형을 비교하기 위해, 선호다양성을 고려하지 않은 일반적인 로짓모형, 다항로짓모형을 먼저 추정하였다.

다항로짓모형의 추정 결과는 <표 24>처럼 나타났으며, 모든 변수의 계수 값이 통계적으로 유의하게 나타났다. 계수의 부호도 통행시간과 통행비용이 직관과 부합하게 음의 방향을 띄고 있다. 승용승합을 기준으로 모든 대안의 대안특성상수 값이 음수가 나타난 것으로 보아, 다른 조건이 모두 동일할 때, 일반적으로 승용승합의 효용이 가장 큰 것을 알 수 있다. 또한 성별과 승용차미보유자인 20대, 고령자, 근무형태 더미변수의 계수 값이 전부 양수로 나타나, 여성이거나 20대일수록 버스의 효용이 증가하고, 고령자, 20대, 전일제 근무자의 경우에는 지하철의 효용이 증가하는 것을 확인할 수 있다. 이러한 결과는 모형을 추정하기 전에 예상한 결과와 추정 결과가 일치하게 나타났다.

하지만, 다항로짓모형의 Rho-squared값이 0.2이하로 나타나 모형의 적합성은 많이 떨어지는 것으로 나타났다.

<표 24> 다항로짓모형 추정 결과

변수	대안	Multinomial Logit Model	
		Coefficient	Standard Error
TT_j	All	-0.8885***	0.3241
TCI_{nj}	All	-2.5312***	0.3521
$ASC_{alternative}$	Passenger	-3.0955***	0.2063
	Bus	-1.9242***	0.1640
	Subway	-0.3300**	0.1439
X_{gender}	Bus	1.1360***	0.2002
X_{noncar}	Bus	0.6371**	0.2930
X_{old}	Subway	0.7116*	0.4173
X_{noncar}	Subway	0.9913***	0.2158
$X_{employee}$	Subway	0.2905**	0.1429
N		1,013	
Log-likelihood		-1038.6860	
Log-likelihood(NULL)		-1111.1930	
ρ^2		0.0653	
$\frac{\rho^2}{\rho^2}$		0.0622	

주: 모든 값은 소수점 다섯째 자리에서 반올림하였음.

***는 99%, **는 95%, *는 90% 신뢰수준에서 통계적으로 유의함을 나타냄.

2) 혼합로짓모형

개인의 선호다양성을 고려한 수단선택 로짓모형을 추정하기 위해 통행시간과 통행비용의 계수분포를 가정하여 혼합로짓모형을 추정한다.

본 연구에서는 일반적으로 가장 많이 사용되는 정규분포와 정규분포의 극단값의 문제점을 보완할 수 있는 삼각분포를 이용한다. 하지만, 정규분포와 일반적인 삼각분포는 부호에 제약이 없어 모든 이용자에게 변수가 동일한 방향으로 영향을 미치는 통행시간과 통행비용의 확률분포로 가정하기에는 문제가 있다. 따라서 본 연구는 모든 이용자에게 동일한 방향으로 효용에 영향을 미치는 것을 반영하기 위해 절단된 정규분포와

절단된 삼각분포로 확률분포를 가정하였다.

통행시간과 통행비용의 계수분포를 절단된 정규분포로 가정한 혼합로짓모형의 추정결과와 절단된 삼각분포로 가정한 혼합로짓모형의 추정결과는 <표 25>와 같다.

<표 25> 혼합로짓모형 추정 결과

변수	대안	Mixed Logit Model			
		Normal_truncated		Triangular_truncated	
		Coefficient	Standard Error	Coefficient	Standard Error
<i>Random Parameters</i>					
TT_j	All	-1.0105***	0.3622	-0.9108***	0.3317
TCI_{nj}	All	-2.2570***	0.3836	-2.5112***	0.3695
<i>Fixed Parameters</i>					
$ASC_{alternative}$	Passenger	-3.1072***	0.2321	-3.1080***	0.2123
	Bus	-1.8455***	0.1639	-1.9128***	0.1646
	Subway	-0.2521*	0.1433	-0.3177**	0.1440
X_{gender}	Bus	1.1181***	0.2029	1.1322***	0.2008
X_{noncar}	Bus	0.7117**	0.3015	0.6507**	0.2947
X_{old}	Subway	0.7195*	0.4304	0.7144*	0.4202
X_{noncar}	Subway	1.0524***	0.2237	1.0024***	0.2173
$X_{employee}$	Subway	0.3052**	0.1465	0.2935**	0.1436
<i>Spread of Parameters</i>					
TT_j		1.0105***	0.3622	0.9108***	0.3317
TCI_{nj}		2.2570***	0.3836	2.5112***	0.3695
N		1,013		1,013	
Log-likelihood		-1044.6495		-1039.7220	
Log-likelihood(NULL)		-1404.3162		-1404.3162	
ρ^2		0.2568		0.2597	
$\overline{\rho^2}$		0.2543		0.2572	

주: 모든 값은 소수점 다섯째 자리에서 반올림하였음.

***는 99%, **는 95%, *는 90% 신뢰수준에서 통계적으로 유의함을 나타냄.

절단된 정규분포로 가정했을 때와 절단된 삼각분포로 가정했을 때 모두 모든 변수의 계수 추정치가 통계적으로 유의한 것으로 나타났다. 통행시간과 통행비용의 계수의 부호도 음의 방향으로 시간이나 비용이 증가할수록 개인의 효용이 감소한다는 직관과 일치한다. 혼합로짓모형의 경우도 다항로짓모형과 동일하게 대안특성상수의 값이 모두 음수로 나타나 다른 조건이 모두 동일할 때, 승용승합의 효용이 가장 큰 것으로 나타났다. 그리고 개인의 사회·경제적 특성 변수에서도 여성이고 20대 일수록 버스의 효용이 증가하고, 고령자이고 20대이며 전일제 근무자일수록 지하철의 효용이 증가하는 것을 확인할 수 있다.

추정된 혼합로짓모형의 계수 부호는 다항로짓모형과 일치하고, 추정치도 확인한 차이가 나는 것은 아니지만, 모형의 적합도를 나타내는 ρ^2 값과 $\overline{\rho^2}$ 값이 약 0.26으로 다항로짓모형에 비해 많이 향상된 것을 알 수 있다. 이를 통해 관측된 자료만을 이용하여 수단선택 모형을 추정할 경우보다 관측되지 않은 개인의 선호다양성까지 고려한 경우, 모형의 적합도가 전반적으로 향상하는 것을 확인할 수 있다.

3) 잠재계층모형

서울시 통근자의 선호다양성을 고려한 수단선택 로짓모형을 추정하기 위해 비슷한 속성을 가지는 표본끼리 계층으로 구분하여 잠재계층모형을 추정하였다.

잠재계층을 나누는 기준으로는 최소 BIC나 최소 AIC의 방법이 사용된다. 최적 계층 수를 결정하기 위해 계층을 1~5개로 나눠서 각각의 BIC와 AIC값을 구하였다. 최소 BIC값은 계층이 2개일 때지만, 최소 AIC값은 계층이 3개일 때로 나타났다. 하지만, AIC의 감소 폭이 계층이 1개에서 2개가 될 때가 계층이 2개에서 3개로 될 때보다 더 크게 나타나 본 연구에서는 최적 계층 수를 2개로 선정하였다.

<표 26> 최적 계층 수 결정 기준

계층 수	Log-likelihood	BIC	AIC	모수	p-value
1	-3127.817	6324.830	6275.634	10	< 0.0001
2	-3061.954	6269.221	6165.907	21	0.001
3	-3044.926	6311.283	6153.853	32	0.117
4	-3034.521	6366.588	6155.041	43	0.515
5	-3031.412	6436.486	6170.823	54	0.114

<표 27> 잠재계층모형 계층 구분 및 계층 특성

(단위 : 명, %)

계층을 구분하는 사회·경제적 특성		표본	Latent Class Model			
			Class 1		Class 2	
			수	비율	수	비율
		1,013	409	40.38	604	59.62
age	20	128	122	29.83	6	0.99
	30	305	172	42.05	133	22.02
	40	314	69	16.87	245	40.56
	50	197	39	9.54	158	26.16
	60	63	6	1.47	57	9.44
	70	6	1	0.24	5	0.83
gender	여	298	292	71.39	6	0.99
	남	715	117	28.61	598	99.01
employee	1(재택근무)	7	4	0.98	3	0.49
	2(전일제)	698	274	66.99	424	70.20
	3(시간제)	86	52	12.71	34	5.63
	4(자영업)	142	33	8.07	109	18.05
	5(기타)	80	46	11.25	34	5.63

따라서 본 연구에서는 잠재계층모형의 계층을 구분하는 기준으로 더 미변수로 사용하는 성별과 연령, 근무형태를 이용하여 구분하였다. 계층을 2개로 구분한 결과는 <표 28>과 같이 나타나며, Class 1은 여성이거나 청년층, Class 2는 남성과 중년과 노년층으로 구분된다. 전체 표본에 남성과 중년층의 표본의 수가 많으므로 Class 2의 비율이 Class 1보다 더 높은 것은 당연한 결과이다. 근무형태의 경우에는 전체 표본에서 전일제가 가장 많은 비중을 차지하고 있는데다가 계층을 두 개로만 구분하기 때문에 근무형태가 뚜렷하게 나뉘지 않아 두 계층 모두 전일제가 차지하는 비율이 높게 나타난 것으로 보인다.

<표 28> 잠재계층모형의 계층 별 수단분담률

(단위 : 통행, %)

		승용승합	승용차동승	버스	지하철	합계
표본	통행량	361	35	128	489	1,013
	비율	35.63	3.46	12.64	48.27	100.00
Class 1	통행량	37	21	87	264	409
	비율	9.05	5.13	21.27	64.55	100.00
Class 2	통행량	324	14	41	225	604
	비율	53.64	2.32	6.79	37.25	100.00

구분된 계층 별 수단분담률을 확인해보면, Class 1의 경우 지하철과 버스의 수단분담률이 높고, Class 2의 경우에는 승용승합과 지하철의 비율이 높은 것으로 나타나, 여성이고 젊은 계층일수록 대중교통을 많이 이용하고, 남성과 30대 이상의 중장년층의 경우 승용차를 많이 이용하고, 정시성의 확보를 위해 지하철도 많이 이용하는 것으로 나타났다.

이렇게 동질성을 띄는 통행자를 동일한 계층으로 구분하고, 각각의 계층끼리는 다양성을 띤다고 가정한 잠재계층모형의 추정결과는 <표 29>와 같다.

<표 29> 잠재계층모형 추정 결과

변수	대안	LCM			
		Class1		Class2	
		Coefficient	Standard Error	Coefficient	Standard Error
TT_j	All	-1.4777***	0.5410	-0.6973*	0.3573
TCI_{nj}	All	-2.6914***	0.7296	-2.9844***	0.3713
$ASC_{alternative}$	Passenger	-1.8786***	0.3577	-3.9149***	0.2319
	Bus	-0.4468	0.4334	-2.5191***	0.1675
	Subway	1.2619***	0.2821	-1.3579***	0.1719
X_{gender}	Bus	0.5605	0.3847	6.0124***	0.2747
X_{noncar}	Bus	0.4832	0.4459	0.1657	0.4423
X_{old}	Subway	19.2722	0.1051	1.1957***	0.3358
X_{noncar}	Subway	0.7592**	0.3787	0.1586	0.2586
$X_{employee}$	Subway	-0.0948	0.2460	0.7406***	0.1653
N		1,013			
Log-likelihood		-1038.7070			
Log-likelihood(NULL)		-1404.3162			
ρ^2		0.2986			
$\overline{\rho^2}$		0.2930			

주: 모든 값은 소수점 다섯째 자리에서 반올림하였음.

***는 99%, **는 95%, *는 90% 신뢰수준에서 통계적으로 유의함을 나타냄.

잠재계층모형의 추정 결과, Class 1과 Class 2 모두 가장 중요한 통행시간과 통행비용의 계수는 직관과 부합하는 음의 부호를 띄며, 그 값이 유의한 것으로 나타났다. 하지만, Class 1과 Class 2의 더미변수는 각 계층을 이루고 있는 특성에 따라 결과가 다르게 나타났다. Class 1의 경우에는 여성과 청년층으로 이루어져있으며, 지하철에 있어서만 20대의 변수가 유의미한 것으로 보아, 청년층은 주로 지하철을 선호하는 것으로 나타났다. Class 2는 주로 남성과 중장년층으로 이루어져있고, 성별과 고령자, 근무형태의 계수추정치이 유의하게 나타난 것으로 보아, 고령자와

전일제 근무자일수록 지하철을 선호하는 것을 알 수 있다.

Class 1의 경우에는 본 연구에서 추정한 다른 로짓모형과는 다른 계수의 부호를 보이는데, 지하철의 상수가 양수로 나타났다. 다른 조건이 동일한 경우 Class 1의 구성원에게는 지하철이 가장 효용이 큰 대안이 될 수 있으며, 이는 자차를 가지고 있지 않은 청년층으로 계층이 이루어져 있기 때문이다. Class 2는 개인의 사회·경제적 특성을 고려하지 않은 경우에는 승용승합의 효용이 가장 큰 것으로 나타났으나, 개인의 사회·경제적 특성을 고려하면, 더미변수의 계수추정치가 모두 양의 부호를 가지고 있어 최대효용수단이 달라진다. 남성이고 중장년층의 계층임을 고려하였을 때, 고령자와 전일제 근무자일수록 지하철이 최대효용수단이 될 가능성이 높음을 확인할 수 있다.

잠재계층모형의 경우 다른 모형에 비해 모든 계수추정치 값이 유의하지는 않지만, Rho-squared 값이 거의 0.3에 가까운 값을 보이는 것을 알 수 있다.

<표 30> 수단선택 로짓모형 추정 결과

변수	대안	MNL	MXL		LCM	
			Normal_truncated	Triangular_truncated	Class1	Class2
TT_j	All	-0.8885***	-1.0105***	-0.9108***	-1.4777***	-0.6973*
TCI_{nj}	All	-2.5312***	-2.2570***	-2.5112***	-2.6914***	-2.9844***
$ASC_{alternative}$	Passenger	-3.0955***	-3.1072***	-3.1080***	-1.8786***	-3.9149***
	Bus	-1.9242***	-1.8455***	-1.9128***	-0.4468	-2.5191***
	Subway	-0.3300**	-0.2521*	-0.3177**	1.2619***	-1.3579***
X_{gender}	Bus	1.1360***	1.1181***	1.1322***	0.5605	6.0124***
X_{noncar}	Bus	0.6371**	0.7117**	0.6507**	0.4832	0.1657
X_{old}	Subway	0.7116*	0.7195*	0.7144*	19.2722	1.1957***
X_{noncar}	Subway	0.9913***	1.0524***	1.0024***	0.7592**	0.1586
$X_{employee}$	Subway	0.2905**	0.3052**	0.2935**	0.0948	0.7406***
N		1,013	1,013	1,013	1,013	
Log-likelihood		-1038.6860	-1044.6495	-1039.7220	-1038.7070	
Log-likelihood(NULL)		-1111.1930	-1404.3162	-1404.3162	-1404.3162	
ρ^2		0.0653	0.2568	0.2597	0.2986	
$\overline{\rho^2}$		0.0622	0.2543	0.2572	0.2930	

주: 모든 값은 소수점 다섯째 자리에서 반올림하였음.

***는 99%, **는 95%, *는 90% 신뢰수준에서 통계적으로 유의함을 나타냄.

제 2 절 모형 간 추정결과의 비교 분석

1) 적합성

(1) 계수의 통계적 유의성

모형의 적합성을 비교하는 방법으로는 Mc Fadden의 ρ^2 값과 수정된 ρ^2 값인 $\overline{\rho^2}$ 을 비교하는 방법이 있다. 앞서서도 설명했듯이 ρ^2 값은 선형 회귀모형의 적합도를 판단하는 R^2 와 동일한 기능을 하며, 일반적으로 0.2에서 0.4의 값을 가지면 유의한 것으로 볼 수 있다.

본 연구에서 추정한 모형의 결과를 살펴보면, 개인의 선호다양성을 고려한 로짓모형인 혼합로짓모형과 잠재계층모형의 경우 ρ^2 와 $\overline{\rho^2}$ 값 모두 0.2이상의 값을 가지고 있으므로 모형의 설명력은 높다고 할 수 있다.

ρ^2 와 $\overline{\rho^2}$ 값은 잠재계층모형이 가장 높은 것으로 나타났고, 이는 개인의 수단선택 로짓모형을 추정할 때, 개인의 선호다양성을 고려하는 것이 설명력이 높은 것을 의미한다.

<표 31> 모형의 적합성 비교

	MNL	MXL		LCM
		Normal_tru	Triangular	
		ncated	_truncated	
N	1,013	1,013	1,013	1,013
Log-likelihood	-1038.6860	-1044.6495	-1039.7220	-1038.7070
ρ^2	0.0653	0.2568	0.2597	0.2986
$\overline{\rho^2}$	0.0622	0.2543	0.2572	0.2930
AIC	2.0705	2.0804	2.0725	1.9922
BIC	2.1190	2.1289	2.1210	2.1088

주: 모든 값은 소수점 다섯째 자리에서 반올림하였음.

(2) 적용 및 확인

추정된 수단선택 로짓모형의 적합성을 판단하기 위해, 추정에 사용되지 않은 표본으로 추정 식에 적용하였다. 추정된 수단선택 로짓모형 식에 적용표본의 통행시간, 분당소득 당 통행비용, 개인의 사회·경제적 특성 값을 대입하여 확인하였다.

최대효용수단의 번호와 실제이용수단의 번호를 비교하였고, 그 결과는 <표 32>와 같다. 승용차동승과 버스 이용자는 최대효용수단이 아님에도 두 수단을 이용하는 것으로 나타났으며, 이는 본 연구에서 확인하지 못한 개인의 다른 선호도가 반영된 것으로 보인다.

<표 32> 최대효용수단과 실제이용수단의 일치 여부

$CM=1$	MNL	MXL		LCM	
		Normal_truncated	Triangular_truncated	Class 1	Class 2
승용승합(1)	65	62	64	0	96
승용차동승(2)	0	0	0	0	0
버스(3)	0	0	0	0	0
지하철(4)	169	175	170	104	47
$\sum_{n=1}^N CM$	234	237	234	104	143
N	434	434	434	175	259
VC	53.92%	54.61%	53.92%	59.43%	55.21%

최대효용수단을 실제로 이용한 통근자가 모든 모형에서 50%이상 나타났기 때문에 모든 추정된 수단선택 로짓모형은 유의미한 것을 알 수 있다. 그 중에서도 잠재계층모형의 Class 1에서 가장 많은 이용자가 최대효용수단을 선택했고, 그 다음 Class 2에서 많은 이용자가 최대효용수단을 선택한 것을 알 수 있다. 이를 통해 잠재계층모형이 가장 개인의 선호도를 잘 반영한 수단선택 로짓모형임을 유추할 수 있다.

다음으로, 적용표본의 실제 수단분담률과 추정분담률을 비교해보았

다. 실제분담률은 승용승합이 35.48%, 승용차동승 3.46%, 버스 12.67%, 지하철 48.39%로 지하철이 가장 높게 나타났다. 모든 수단선택 로짓모형에서 승용차동승과 버스는 최대효용수단이 아닌 것으로 나타났다. 그 이유는 아무래도 통근통행을 대상으로 하고, 실제 자료에서도 승용승합과 지하철의 수단분담률이 높기 때문이다. 또한 대안특성상수가 잠재계층모형 Class 1의 지하철 상수를 제외하고는 모두 음의 값을 띄고 있으므로, 모든 조건이 동일한 경우에는 승용승합이 최대효용수단이고, 더미변수를 포함한 경우는 지하철 효용함수 식에 고정자, 20대, 근무형태의 3개의 더미변수가 포함되고, 각각의 계수추정치가 양수의 값으로 나타나기 때문에 지하철이 최대효용함수가 될 가능성이 높기 때문에 이러한 결과로 나타났을 가능성이 있다. 실제분담률과 가장 비슷한 추정분담률을 보이는 모형은 잠재계층모형이므로, 잠재계층모형이 가장 설명력이 높다.

<표 33> 적용 표본의 실제분담률과 추정분담률

(단위 : 통행, %)

분담률		승용승합	승용차동승	버스	지하철	합계	
실제분담률	통행량	154	15	55	210	434	
	비율	35.48	3.46	12.67	48.39	100.00	
MNL	통행량	120	0	0	294	434	
추정분담률	비율	27.65	0.00	0.00	67.74	100.00	
MXL(N)	통행량	110	0	0	324	434	
추정분담률	비율	25.35	0.00	0.00	74.65	100.00	
MXL(T)	통행량	117	0	0	317	434	
추정분담률	비율	26.96	0.00	0.00	73.04	100.00	
추정분담률	LCM	통행량	175	0	0	259	434
		비율	40.32	0.00	0.00	59.68	100.00
	Class1	통행량	0	0	0	175	175
		비율	0.00	0.00	0.00	100.00	100.00
	Class2	통행량	175	0	0	84	259
		비율	67.57	0.00	0.00	32.43	100.00

2) 시간가치 및 탄력성

본 연구에서는 제 3장에서 제시한 모형의 추정결과 비교 지표로 각 모형별 통행시간가치와 통행시간과 통행비용에 관한 수요 탄력성을 구하고, 각각을 비교하고자 한다.

통행시간가치는 통행시간이 감소할 때 개인이 지불하고자 하는 금전적 가치를 의미하며, 추정된 통행시간과 통행비용 계수 비에 월평균소득을 곱한 값으로 구할 수 있다. 표본의 월평균 분당소득은 375.62원으로, 추정된 다항로짓모형의 계수 비를 이용하여 통행시간가치를 산정한다. 혼합로짓모형의 경우, 개인 별 계수를 추정할 수 있으므로 모집단의 분포를 추정하고, 혼합로짓모형의 통행시간가치 산정 식에 의해 분포별 통행시간가치를 산정하였다. 절단된 삼각분포의 통행시간가치가 절단된 정규분포의 통행시간가치보다 높게 나타난 것은, 정규분포의 경우에는 계수추정치에 0에 더 많이 몰려있지만, 삼각분포의 경우에는 0점에서 최댓값까지 골고루 퍼져있어 정규분포보다 큰 값에 더 많이 분포하고 있어 통행시간가치가 높게 나타난 것을 알 수 있다.

<표 34> 통행시간가치

(단위: 원/시간)

	Multinomial Logit Model	Mixed Logit Model		Latent Class Model	
		Normal_truncated	Triangular_truncated	Class1	Class2
VOT	7,911	6,808	8,980	11,572	5,497

주: 모든 값은 소수 첫째자리에서 반올림한 값임.

잠재계층모형의 경우에는 계층별로 소득수준이 다르기 때문에, 각각의 계층의 월평균소득을 계수 비에 곱한 값을 사용하였다. Class 1의 월평균 분당소득은 351.27원이고, Class 2의 월평균 분당소득은 392.11원으로 나타났다. 따라서 각 계층별 월평균 분당소득과 통행시간, 통행비용의

계수 비를 이용하여 통행시간가치를 구하면, Class 1은 11,572원, Class 2는 5,497원으로 나타났다. 일반적으로 평균소득이 높을수록 시간 당 소득이 높기 때문에 시간가치가 더 높을 것으로 예상되지만, 월평균 분당 소득의 차이가 그리 크지 않으며, Class 1과 Class 2에서 소득을 제외한 다른 개인의 사회·경제적 특성이 시간가치에 더 많은 영향을 주기 때문에 직관과는 다른 결과가 나온 것으로 보인다.

선호다양성을 고려한 수단선택 로짓모형에 관한 선행연구 중 모형 간 비교방법으로 시간가치를 산정하고, 그 값을 비교한 연구는 <표 37>과 같다.

<표 35> 선행연구 중 시간가치를 비교한 연구

선행연구	자료 형태	모형	모형 비교방법		시간가치가 가장 높은 모형
			시간가치	탄력성	
정진혁 외(2007)	가구통행 실태조사	MNL/MXL	○		△
이준규 외(2015)	SP	MNL/MXL	○		MXL
조신형 외(2017)	SP	MNL/MXL	○		MXL
Greene and Hensher(2003)	SP	MNL/MXL/LCM	○	○	MXL
Viton(2004)	RP	MNL/MXL	○		MXL
Shen(2009)	SP	MNL/MXL/LCM	○	○	△
Greene and Hensher(2010)	SP	MNL/MXL	○	○	△
Zito and Salvo(2012)	SP	MNL/LCM	○		LCM
Leon and Miguel(2017)	SP	CL/MXL	○		△
Tirachini et al.(2017)	SP	MNL/MXL/LCM	○		MXL

주: MNL(Multinomial Model), CL(Conditional Logit Model), MXL(Mixed Logit Model), LCM(Latent Class Model), LC-MXL(Latent Class-Mixed Logit Model).

△는 어떤 특정한 모형의 시간가치가 가장 높다고 말할 수 없는 경우.

선행연구에서도 대체로 선호다양성을 고려했을 때의 모형이 시간가치가 가장 높은 모형으로 나타났다. 또한, 통행시간을 대기시간, 주행시간, 접근시간 등으로 세분화하여 각각의 통행시간가치를 산정한 연구한 경우에는 세분화된 통행시간 변수별로 통행시간가치가 가장 높은 모형이 달라지고, 특히 잠재계층모형의 경우에는 계층별로 구분된 개인의 통행시간가치의 영향을 받기 때문에 어떤 모형이 가장 시간가치가 높은 모형이라고 단정 짓지 못하는 것을 확인할 수 있다.

통행시간과 통행비용에 관한 수요 탄력성은 통행시간과 통행비용의 평균값으로 각각의 탄력성이 계산되었다. 모형 간 통행시간과 통행비용에 관한 수요 탄력성을 비교해보면, 통행시간과 통행비용 모두 잠재계층모형의 탄력성의 절댓값이 가장 크게 나타났다. 이는 통행시간과 통행비용이 1% 증가하거나 감소함에 따라 이용자가 가장 민감하게 반응하는 모형이 잠재계층모형이라는 것이며, 잠재계층모형이 이러한 개인의 선호다양성을 잘 반영하는 모형이라고 할 수 있다.

<표 36> 로짓모형 별 통행시간과 통행비용에 대한 수요 탄력성

	Multinomial Logit Model	Mixed Logit Model		Latent Class Model	
		Normal_trun cated	Triangular_ truncated	Class1	Class2
통행시간탄력성					
승용승합	-0.8902	-0.9547	-0.9005	-2.1983	-0.5265
승용차동승	-1.3186	-1.5517	-1.3582	-2.2891	-1.0834
버스	-1.2804	-1.3635	-1.2980	-2.0652	-1.1009
지하철	-0.7606	-0.8116	-0.7700	-0.9238	-0.7490
통행비용탄력성					
승용승합	-1.1319	-0.7225	-1.0434	-1.4683	-1.2230
승용차동승	-0.8665	-0.9214	-0.8892	-0.8271	-1.2384
버스	-1.0198	-0.9864	-1.0259	-0.8744	-1.5558
지하철	-0.6404	-0.5357	-0.6240	-0.4253	-1.1057

일반적으로 소득이 낮거나 시간가치가 높을수록 시간이나 요금 변화에 더 민감하다. 본 연구에서도 일반적인 상식과 부합하게 월평균 소득이 낮고 시간가치가 가장 높은 Class 1이 통행시간에 있어서 가장 민감한 것으로 나타났다. 하지만, 통행비용에 있어서는 승용승합에서만 Class 1이 가장 민감한 것으로 나타났고, 나머지 수단에 대해서는 Class 2가 요금 변화에 가장 민감한 것으로 나타났다.

다항로짓모형과 절단된 삼각분포로 가정한 혼합로짓모형의 승용승합을 제외하고는 모든 모형에서 통행시간탄력성이 통행비용탄력성보다 높게 나타났다. 이는 통근자는 비용이 증가하는 것보다 시간이 증가하는 것에 더 민감하게 반응하는 것을 의미한다. 하지만, 승용승합의 경우에는 통행비용탄력성이 더 높은 모형이 존재했는데, 그 이유는 승용차의 경우 대중교통 요금에 비해 유류비, 주차비 등이 차지하는 비용이 크기 때문이다. 이러한 비용이 증가할 경우 승용차를 이용하던 사람들이 다른 수단으로의 전환 가능성이 높은 것으로 보인다.

또한 모든 모형에서 통행시간과 통행비용의 탄력성이 4개의 대안 중 지하철에 대한 수요 탄력성이 가장 덜 민감한 것으로 나타났는데, 이는 통근통행이라는 특성과 오전 침두시라는 점을 고려했을 때 정시성을 보장하는 지하철에 대해서는 이용자들이 쉽게 수단을 변경하지 않는 것을 의미한다.

제 6 장 결론

제 1 절 결론

기술의 발전과 사회가 변화함에 따라 과거에 비해 개인의 특성이 다양해지고, 이용 가능한 교통수단의 종류가 증가하였다. 이러한 다양성으로 인해 기존에 수단선택 로짓모형으로 사용되던 이항로짓모형이나 다항로짓모형 이외의 개인의 다양성을 반영한 발전된 모형의 필요성이 대두되고 있다.

개인의 선호다양성을 고려한 수단선택 로짓모형에 관한 연구는 국내외적으로 다양하게 진행되고 있지만, 그 역사는 그리 길지 않으며 본 연구에서 다루고자 하는 다항로짓모형, 혼합로짓모형, 잠재계층모형 세 모형의 비교연구는 아직까지 이루어지지 않았다.

대부분의 선행 연구는 특정 상황에 대한 설문조사를 기반으로 한 연구이다. 설문조사를 기반으로 할 경우 그 상황에서 응답자의 실제 행동과는 차이가 발생할 수 있다는 문제점이 있다. 따라서 본 연구에서는 서울시에서 개인의 실제 통행을 기반으로 하여 5년마다 실시하는 가구통행실태조사 자료를 이용하여, 서울시 통근자를 대상으로 개인의 선호다양성을 고려한 수단선택 로짓모형을 추정하고자 한다.

수단선택 모형 추정 시 이용하는 변수에는 자료에서 관측 가능한 통행시간, 통행비용 변수뿐만 아니라 사회·경제적 특성변수를 포함시켜, 개인의 다양성을 고려하였고, 소득에 따라 동일한 통행요금이라도 개인이 체감하는 비용은 다르기 때문에 월평균 분당소득 대비 통행비용을 이용하였다. 이용자가 이용할 수 있는 수단은 승용승합, 승용차동승, 버스, 지

하철 4가지로 한정하였고, 승용차동승을 고려하는 점에서 공유시장이 점차 확대되고 있는 현재의 모습을 반영하였다는 점에서 선행연구와의 차별성을 보인다.

개인의 사회·경제적 특성에 따라 교통수단에 대한 개인의 효용이 다를 것을 고려하여, 수단 별로 더미변수를 다르게 하여 효용함수 식을 산정하고 모형을 추정하였다. 추정된 모형의 적합성을 확인하기 위해 ρ^2 값을 비교하였고, 그 결과 잠재계층모형의 ρ^2 값이 0.2986으로 가장 높게 나타나 설명력이 가장 높은 것으로 확인되었다. 또한 추정된 모형이 다른 표본에도 동일하게 적용이 되는지를 확인하기 위해 추정에 사용된 표본 외의 검증 표본을 이용하여 확인하는 과정을 거쳤다. 그 결과 모든 모형에서 50%이상의 통행자가 실제로 효용이 가장 높은 수단을 이용한 것으로 나타났으며, 실제분담률과 추정분담률을 비교하여, 잠재계층모형에서 가장 비슷한 수단분담률을 보이는 것을 확인하였다.

또한 어떤 모형이 가장 개인의 선호다양성을 나타내는지 확인하기 위해 통행시간 한 단위당 개인의 통행시간가치와 통행시간과 통행비용의 증감에 따른 수요의 탄력성을 구하였다. 그 결과, 통행시간가치는 모형에 따라 다르게 나타났으며, 혼합로지트모형의 경우 절단된 정규분포보다 절단된 삼각분포에서 통행시간가치가 더 크게 나타났다. 삼각분포는 정규분포에 비해 극한값을 덜 가지며 값이 골고루 퍼져있어 상대적으로 큰 값이 더 많이 분포하여 통행시간가치가 높게 나타난 것으로 보인다. 잠재계층모형의 경우에는 계층에 속하는 개인의 특성에 따라 통행시간가치가 다르게 나타나 어떠한 기준으로 계층을 구분하느냐에 따라 그 값이 달라짐을 알 수 있었다.

통행시간과 통행비용의 수요탄력성은 모두 잠재계층모형일 때 가장 민감한 것으로 나타나, 잠재계층모형이 가장 개인의 선호다양성을 잘 반영하는 모형임을 알 수 있었다. 통행시간과 통행비용의 수단별 탄력성을 비교해보면, 모든 모형에서 지하철의 탄력성이 가장 덜 민감한 것으로

나타나고, 대부분의 모형과 수단에서 시간에 대한 탄력성이 비용에 대한 탄력성 보다 더 크게 나타났다. 이는 오전 첨두시의 통근자만을 대상으로 했기 때문에 통근통행의 특성 상 정시성이 수단선택에 중요한 결정요인으로 시간의 변화에 통근자가 더 민감하고, 지하철의 통행시간이나 통행비용이 증가하더라도 다른 수단에 비해 정시성이 보장되는 수단이므로 수요가 덜 감소하는 것으로 나타났다.

본 연구는 수단선택 로짓모형을 개인의 선호다양성을 고려한 다양한 모형으로 추정하고 이를 해석하여 잠재계층모형이 가장 설명력이 높은 모형임을 도출함으로써, 점차 세분화되는 개인의 특성을 잘 반영할 수 있는 모형을 제안하는 것에 의의가 있다. 따라서 현재 이용하고 있는 일반적인 수단선택 로짓모형 대신 개인의 선호다양성을 반영한 모형으로 개편이 필요하다.

제 2 절 한계 및 향후 연구과제

개인의 선호다양성을 고려한 수단선택 로짓모형을 추정하기 위해서 일반적으로 SP조사를 이용하여 연구를 진행해왔다. 본 연구에서는 SP조사의 한계점을 개선하기 위해 가구통행실태조사 자료의 실제 통행과 선택한 수단 자료를 이용하여 개인의 선호다양성을 고려한 수단선택 로짓모형을 추정하였지만, 몇 가지 한계점을 가진다.

가구통행실태조사 자료는 선택한 수단에 대한 자료는 제공하지만, 선택하지 않은 수단에 대한 자료는 제공하지 않는다. 따라서 본 연구에서는 네이버 지도에서 제공하는 최단 소요시간의 수단을 선택한다고 가정하였고, 행정동 정보를 이용하여 행정동 주민센터 간 이동을 기준으로 하였다. 하지만, 실제로 통행자의 거주지나 직장이 행정동 주민센터와 일치하지 않기 때문에 거주지와 직장의 위치에 따라 경로나 수단이 다를 수 있다.

두 번째로 가구통행실태조사 자료에서 제공하는 통행자의 실제 통행 시간은 응답자의 기억에 의존하여 기재된 시간이므로, 정확한 자료가 아니라 대략적인 시간인 경우가 많아 통행시간 자료의 정확성이 낮다. 이러한 문제점을 해결하기 위해 좀 더 정확한 시간과 출도착 지점의 자료를 확보하는 것이 중요하지만, 스마트카드 자료의 경우에는 개인 프라이버시의 문제로 인해 개인정보를 제외한 자료만을 제공하는 한계가 있다.

마지막으로 본 연구는 연구에서 사용한 자료가 2010년 가구통행실태조사 자료이기 때문에 2010년 기준으로 개통되지 않은 지하철 노선에 관한 자료는 대안이 4개가 될 수 없다고 생각하여 표본에서 제외하였다. 하지만, 지하철을 이용한 경우에는 제외된 표본이 거의 없기 때문에 표본의 수단분담률이 일반적인 수단분담률에 비해 지하철의 수단분담률이 높게 측정되었다.

그럼에도 불구하고 실제 통행을 기반으로 하는 가구통행실태조사 자료를 이용하여 수단선택 로짓모형을 추정하였으므로 SP자료로 추정한 연구에 비해 현실을 더 잘 반영하는 것은 틀림없다. 본 연구에서는 자료 제공의 문제로 이용하지 못하였지만, 최신의 자료일수록 현재와 가장 비슷한 통행특성을 반영할 것이므로 향후 제공되는 최신 자료로 수단선택 로짓모형을 추정한다면 현실을 더 잘 반영할 수 있을 것이며, 스마트카드와 같은 자료를 함께 이용할 수 있다면 더 발전된 결과가 도출될 것이라고 예상된다.

참 고 문 헌

- 강수철·남승용·김만배(2009), “직장인의 차량보유 결정요인 및 통근수단 선택행태 분석”, 「한국정책과학학회보」, 13(1): 287-302.
- 고용노동부(2010), 사업체 임금근로시간조사 보고서.
- 구자춘(2013), 도시립의 속성에 대한 도시민의 선호체계-선택실험법 적용을 중심으로, 서울대학교 대학원 박사학위논문.
- 국가교통DB센터, 2006년 광역권 여객통행실태조사, <https://www.ktddb.go.kr/www/contents.do?key=65>, 2006.
- 권기현·전명진(2014), “생활시간 데이터를 활용한 수도권 맞벌이 가구의 성별 생활시간 및 이동수단 차이 분석”, 「지역연구」, 30(3): 109-129.
- 김강수·조혜진(2006), 「SP 조사설계 및 분석방법론」, 서울: 보성각.
- 배윤경·정진혁·김형진(2010), “잠재계층분석에 따른 수단선택모형비교분석”, 「대한교통학회지」, 28(3): 99-107.
- _____. _____. _____. (2011), “환경특성을 반영한 잠재계층수단선택모형 개발”, 「지역연구」, 27(2): 99-111.
- 수도권교통본부(2012), 여객 기종점통행량(O/D) 전수화 및 장래수요예측 공동조사.
- 유병국(2012), “선택실험법 자료에서의 선호이질성 분석을 위한 혼합로짓 모형 및 잠재계층모형의 활용”, 「지원·환경경제연구」, 21(4): 921-945.
- 윤대식(2001), 「교통수요분석」, 서울: 박영사.
- 이성우·민성희·박지영·윤성도(2005), 「로짓·프라빗모형 응용」, 서울: 박영사.

이혜승·이혜승(2009), “서울시 대중교통체계 개편 이후 통근 교통수단 선택의 차별적 변화”, 「대한지리학회지」, 44(3): 323-338.

이희연·노승철(2012), 「고급통계분석론 - 이론과 실습」, 경기: 문우사.

전은하·이성우(2014), “수도권 교차통근자의 통행수단 선택 분석 : 1995-2000”, 「서울도시연구」, 8(4): 107-125.

정진혁·정진욱(2007), “혼합로짓모형을 이용한 통행시간가치 신뢰구간 추정”, 「지역연구」, 23(3): 91-109.

진중헌(1994), “직장과 주거지의 입지적 상호의존관계에 관한 연구 - 성별 차이를 중심으로 -”, 「지리학회지」, 24: 175-196.

최영은(2017), 유료도로 선택행태를 이용한 통행시간 신뢰성의 가치 추정, 서울대학교 대학원 박사학위논문.

추상호(2012), “활동요소가 통행수단선택에 미치는 영향 분석 : 선택적 활동을 중심으로”, 「국토연구」, 74: 163-173.

한국개발연구원(2008), 도로·철도 부문 사업의 예비타당성조사 표준지침 수정·보완 연구(제5판).

_____ (2013), 교통부문 예비타당성조사 쟁점 연구.

한국주택학회, HUG Brownbag Seminar on Housing Research - 로짓모형을 활용한 주거이동분석, http://www.kahps.org/board/board_view.html?ps_boid=93&ps_db=academic, 2016.

Ben-Akiva M. and Lerman, S. R.(1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, Cambridge and London: The MIT Press.

Boxwall, P. C. and Adamowicz, W. L.(1999), *Understanding Heterogeneous Preferences in Random Utility Models: The Use of*

Latent Class Analysis, University of Alberta Edmonton.

_____ and _____(2002), "Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach", *Environmental and Resource Economics*, 23(4): 421-446.

Bujosa, A., Riera, A. and Hicks, R. L.(2010), "Combining Discrete and Continuous Representations of Preference Heterogeneity: A Latent Class Approach", *Environmental and Resource Economics*, 47(4): 477-493.

David, K.(Ed.)(2004), *The SAGE Handbook of Quantitative Methodology for the Social Sciences*, California: Sage Publications Inc.

Greene, W. H. and Hensher, D. A.(2003), "A Latent Class Model for Discrete Choice Analysis : Contrasts with Mixed Logit", *Transportation Research Part B*, 37: 681-698.

_____ and _____(2010), "Does Scale Heterogeneity across Individuals Matter? An Empirical Assessment of Alternative Logit Models", *Transportation*, 37: 413-428.

Hensher, D. A. and Greene, W. H(2001), *The Mixed Logit Model: The State of Practice and Warnings for the Unwary*, University of Sydney.

_____, Rose, J. M. and Greene, W. H.(2015), *Applied Choice Analysis (2nd ed.)*, Cambridge: Cambridge University Press.

Hess, S., Bierlaire, M. and Polak, J. W.(2005), "Estimation of Value of Travel-time Savings Using Mixed Logit Models", *Transportation Research Part A: Policy and Practice*, 39(2): 221-236.

- Lee, B. J., Fujiwara, A., Zhang, J. and Sugie, Y.(2003), "Analysis of Mode Choice Behaviours based on Latent Class Models", *10th International Conference on Travel Behaviour Research*, 10(15).
- Leon, G. and Miguel, E.(2017), "Risky Transportation Choices and the Value of a Statistical Life", *American Economic Journal: Applied Economics*, 9(1): 202-228.
- Louviere, J. J., Hensher, D. A., Swait, J. D. and Adamowicz, W. (2000), *Stated Choice Methods: Analysis and Applications*, NewYork: Cambridge University Press.
- Shen, J.(2009), "Latent Class Model or Mixed Logit Model? A Comparison by Transport Mode Choice Data", *Applied Economics*, 41(22): 2915-2924.
- Tirachini, A., Hurtubia, R., Dekker, T. and Daziano, R. A.(2017), "Estimation of Crowding Discomfort in Public Transport: Results from Santiago de Chile", *Transportation Research Part A: Policy and Practice*, 103: 311-326.
- Train, K. E.(Ed.)(2009), *Discrete Choice Methods with Simulation*, NewYork: Cambridge University Press.
- Viton, P. A.(2004), "Will Mixed Logit Change Urban Transport Policies?", *Journal of Transport Economics and Policy*, 38(3): 403-423.
- Zito, P. and Salvo, G.(2012), "Latent Class Approach to Estimate the Willingness to Pay for Transit User Information", *Journal of Transportation Technologies*, 2(03): 193-203.

<부록 1-1> 수단별 개인의 사회·경제적 특성(추정 표본)

(단위 : 명)

특성		승용승합	승용동승	버스	지하철	합계
합계		361	35	128	489	1,013
성별	여성	38	15	67	178	298
	남성	323	20	61	311	715
연령	20대	12	4	22	90	128
	30대	88	10	38	169	305
	40대	147	11	36	120	314
	50대	95	4	21	77	197
	60대	18	5	11	29	63
	70대 이상	1	1	0	4	6
운전면허	유	361	30	95	397	883
	무	0	5	33	92	130
근무형태	재택근무	1	2	0	4	7
	전일제	231	24	88	355	698
	시간제	21	3	15	47	86
	자영업	87	4	10	41	142
	기타	21	2	15	42	80
직업종류	전문직	60	10	22	63	155
	서비스업	30	2	29	67	128
	판매업	60	4	12	49	125
	사무직	148	14	51	216	429
	기능노무	20	4	6	36	66
	기타	43	1	8	58	110
소득	저	38	3	20	103	164
	중	237	21	86	285	629
	고	86	11	22	101	220

<부록 1-2> 개인의 사회·경제적 특성별 수단분담률(추정 표본)

(단위 : %)

특성		승용승합	승용동승	버스	지하철	합계
성별	여성	12.75	5.03	22.48	59.73	100.00
	남성	45.17	2.80	8.53	43.50	100.00
연령	20대	9.38	3.12	17.19	70.31	100.00
	30대	28.85	3.28	12.46	55.41	100.00
	40대	46.82	3.50	11.46	38.22	100.00
	50대	48.22	2.03	10.66	39.09	100.00
	60대	28.57	7.94	17.46	46.03	100.00
	70대 이상	16.67	16.67	0.00	66.67	100.00
운전면허	유	40.88	3.40	10.76	44.96	100.00
	무	0.00	3.85	25.38	70.77	100.00
근무형태	재택근무	14.29	28.57	0.00	57.14	100.00
	전일제	33.09	3.44	12.61	50.86	100.00
	시간제	24.42	3.49	17.44	54.65	100.00
	자영업	61.27	2.82	7.04	28.87	100.00
	기타	26.25	2.50	18.75	52.50	100.00
직업종류	전문직	38.71	6.45	14.19	40.65	100.00
	서비스업	23.44	1.56	22.66	52.34	100.00
	판매업	48.00	3.20	9.60	39.20	100.00
	사무직	34.50	3.26	11.89	50.35	100.00
	기능노무	30.30	6.06	9.09	54.55	100.00
	기타	39.09	0.91	7.27	52.73	100.00
소득	저	23.17	1.83	12.20	62.80	100.00
	중	37.68	3.34	13.67	45.31	100.00
	고	39.09	5.00	10.00	45.91	100.00

<부록 1-3> 수단별 개인의 사회·경제적 특성 비율(추정 표본)

(단위 : %)

특성		승용승합	승용동승	버스	지하철	합계
합계		100.00	100.00	100.00	100.00	100.00
성별	여성	10.53	42.86	52.34	36.40	29.42
	남성	89.47	57.14	47.66	63.60	70.58
연령	20대	3.32	11.43	17.19	18.40	12.63
	30대	24.37	28.57	29.69	34.56	30.11
	40대	40.72	31.43	28.13	24.54	31.00
	50대	26.32	11.43	16.41	15.75	19.45
	60대	4.99	14.29	8.59	5.93	6.22
	70대 이상	0.28	2.86	0.00	0.82	0.59
운전면허	유	100.00	85.71	74.22	81.19	87.17
	무	0.00	14.29	25.78	18.81	12.83
근무형태	재택근무	0.28	5.71	0.00	0.82	0.69
	전일제	63.99	68.57	68.75	72.60	68.90
	시간제	5.82	8.57	11.72	9.61	8.49
	자영업	24.10	11.43	7.81	8.38	14.02
	기타	5.82	5.71	11.72	8.59	7.90
직업종류	전문직	16.62	28.57	17.19	12.88	15.30
	서비스업	8.31	5.71	22.66	13.70	12.64
	판매업	16.62	11.43	9.38	10.02	12.34
	사무직	41.00	40.00	39.84	44.17	42.35
	기능노무	5.54	11.43	4.69	7.36	6.52
	기타	11.91	2.86	6.25	11.86	10.86
소득	저	10.53	8.57	15.63	21.06	16.19
	중	65.65	60.00	67.19	58.28	62.09
	고	23.82	31.43	17.19	20.65	21.72

<부록 2-1> 수단별 개인의 사회·경제적 특성(적용 표본)

(단위 : 명)

특성		승용승합	승용동승	버스	지하철	합계
합계		154	15	55	210	434
성별	여성	16	7	28	76	127
	남성	138	8	27	134	307
연령	20대	8	3	14	56	81
	30대	40	4	11	44	99
	40대	55	5	22	65	147
	50대	43	2	7	32	84
	60대	8	1	1	9	19
	70대 이상	0	0	0	4	4
운전면허	유	154	12	43	172	381
	무	0	3	12	38	53
근무형태	재택근무	0	0	0	2	2
	전일제	103	10	39	154	306
	시간제	10	1	7	14	32
	자영업	30	3	3	21	57
	기타	11	1	6	19	37
직업종류	전문직	25	2	8	38	73
	서비스업	19	1	5	19	44
	판매업	24	3	7	18	52
	사무직	58	7	22	89	176
	기능노무	13	0	4	16	33
	기타	15	2	9	30	56
소득	저	17	1	10	40	68
	중	98	10	38	131	277
	고	39	4	7	39	89

<부록 2-2> 개인의 사회·경제적 특성별 수단분담률(적용 표본)

(단위 : %)

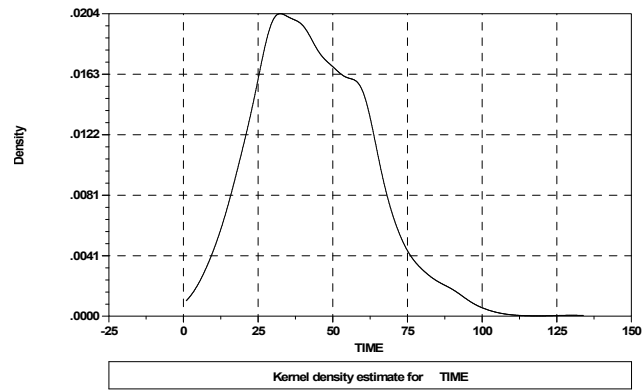
특성		승용승합	승용동승	버스	지하철	합계
성별	여성	12.60	5.51	22.05	59.84	100.00
	남성	44.95	2.61	8.79	43.65	100.00
연령	20대	9.88	3.70	17.28	69.14	100.00
	30대	40.40	4.04	11.11	44.44	100.00
	40대	37.41	3.40	14.97	44.22	100.00
	50대	51.19	2.38	8.33	38.10	100.00
	60대	42.11	5.26	5.26	47.37	100.00
	70대 이상	0.00	0.00	0.00	100.00	100.00
운전면허	유	40.42	3.15	11.29	45.14	100.00
	무	0.00	5.66	22.64	71.70	100.00
근무형태	재택근무	0.00	0.00	0.00	100.00	100.00
	전일제	33.66	3.27	12.75	50.33	100.00
	시간제	31.25	3.13	21.88	43.75	100.00
	자영업	52.63	5.26	5.26	36.84	100.00
	기타	29.73	2.70	16.22	51.35	100.00
직업종류	전문직	34.25	2.74	10.96	52.05	100.00
	서비스업	43.18	2.27	11.36	43.18	100.00
	판매업	46.15	5.77	13.46	34.62	100.00
	사무직	32.95	3.98	12.50	50.57	100.00
	기능노무	39.39	0.00	12.12	48.48	100.00
	기타	26.79	3.57	16.07	53.57	100.00
소득	저	25.00	1.47	14.71	58.82	100.00
	중	35.38	3.61	13.72	47.29	100.00
	고	43.82	4.49	7.87	43.82	100.00

<부록 2-3> 수단별 개인의 사회·경제적 특성 비율(적용 표본)

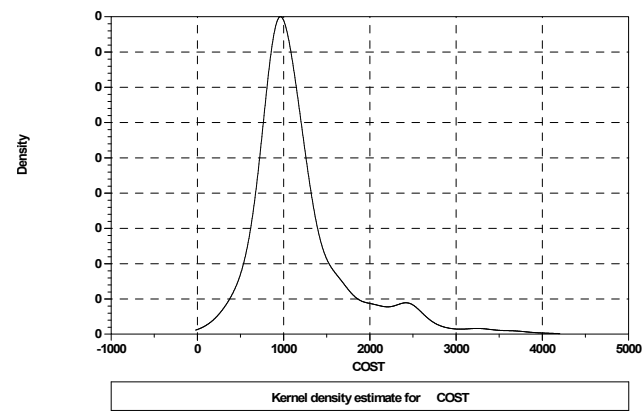
(단위 : %)

특성		승용승합	승용동승	버스	지하철	합계
합계		100.00	100.00	100.00	100.00	100.00
성별	여성	10.39	46.67	50.91	36.19	29.26
	남성	89.61	53.33	49.09	63.81	70.74
연령	20대	5.19	20.00	25.45	26.67	18.66
	30대	25.97	26.67	20.00	20.95	22.81
	40대	35.71	33.33	40.00	30.95	33.87
	50대	27.92	13.33	12.73	15.24	19.35
	60대	5.19	6.67	1.82	4.29	4.38
	70대 이상	0.00	0.00	0.00	1.90	0.92
운전면허	유	100.00	80.00	78.18	81.90	87.79
	무	0.00	20.00	21.82	18.10	12.21
근무형태	재택근무	0.00	0.00	0.00	0.95	0.46
	전일제	66.88	66.67	70.91	73.33	70.51
	시간제	6.49	6.67	12.73	6.67	7.37
	자영업	19.48	20.00	5.45	10.00	13.13
	기타	7.14	6.67	10.91	9.05	8.53
직업종류	전문직	16.23	13.33	14.55	18.10	16.82
	서비스업	12.34	6.67	9.09	9.05	10.14
	판매업	15.58	20.00	12.73	8.57	11.98
	사무직	37.66	46.67	40.00	42.38	40.55
	기능노무	8.44	0.00	7.27	7.62	7.60
	기타	9.74	13.33	16.36	14.29	12.90
소득	저	11.04	6.67	18.18	19.05	15.67
	중	63.64	66.67	69.09	62.38	63.82
	고	25.32	26.67	12.73	18.57	20.51

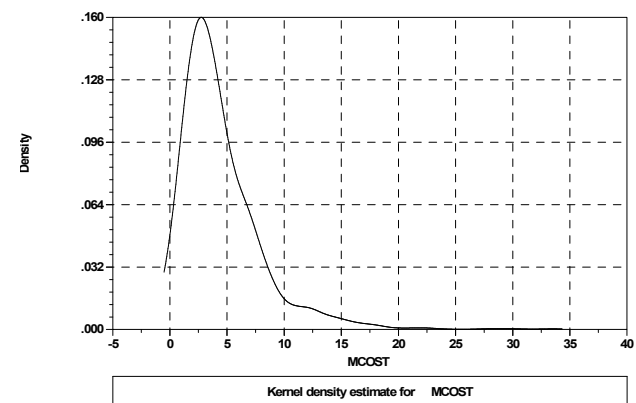
<부록 3-1> 승용차보유가구 통행시간분포



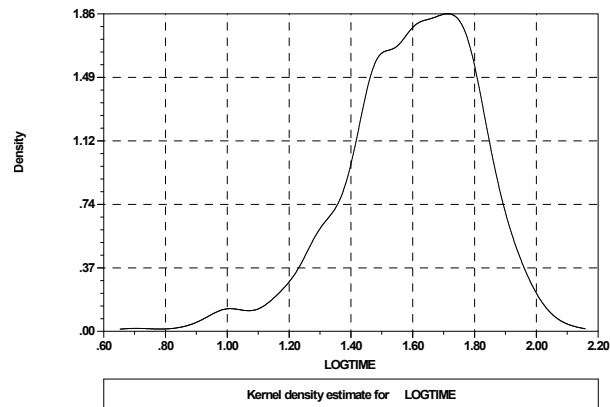
<부록 3-2> 승용차보유가구 통행비용분포



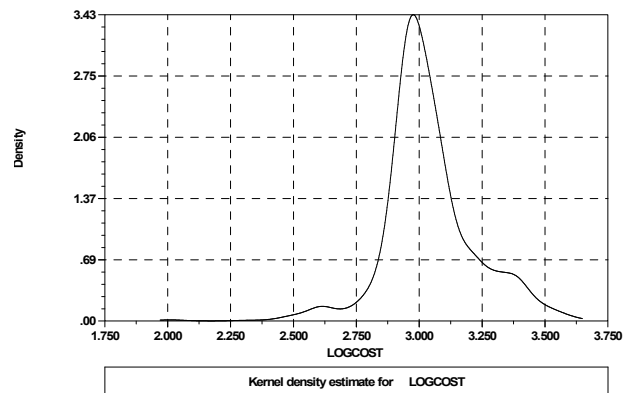
<부록 3-3> 소득대비 통행비용 분포



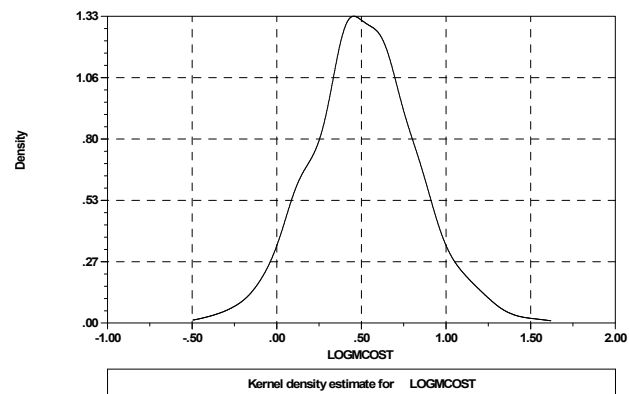
<부록 3-4> 로그통행시간 분포



<부록 3-5> 로그통행비용 분포



<부록 3-6> 로그 소득대비 통행비용분포



Abstract

A Comparative Analysis of
Logit Mode Choice Models Examining
Estimation and Suitability

Minju Kim

Transportation Studies

Department of Environmental Planning

The Graduate School of Environmental Studies

Seoul National University

Following rapid technological development and societal changes, the characteristics of individuals have become more diverse, and available modes of transportation have increased. As a result of this diversity, there is a growing need for an improved model that reflects the diversity of individuals other than the binomial and multinomial logit models currently used as mode choice logit models.

While research on mode choice logit models that consider the diversity of individual preferences is taking place across a variety of studies both domestically and abroad, investigations on this topic have only occurred in recent years, and a comparative analysis of the Multinomial Logit Model (MNL), Mixed Logit Model (MXL), and Latent Class Model (LCM), like the one presented in this study, has

yet to be conducted.

The majority of prior studies are based on surveys about particular situations. However, a problem with basing research on surveys is that respondents' actual behaviors in a situation may differ from those indicated by their responses. Therefore, this study aims to estimate a mode choice logit model that reflects the diversity of individual preferences among Seoul's commuters by using data from the Household Travel Diary Survey, which tracks individuals' actual travel behavior and is conducted every five years in the city of Seoul.

The variables used when estimating mode choice logit models consider the diversity of individuals by including not only variables such as travel time and travel expenses that can be observed in the data but also socio-economic characteristics variables. Because travel costs felt by individuals will differ even with the same travel cost rate due to differences in income, the cost of travel divided by monthly average income per minute is used. The modes available for users are limited to four types: private vehicles driven by the individual in question, private vehicles in which the individual is a passenger but not the driver, buses, and the subway. This study differentiates itself from prior studies by considering that private car passengers may not be the driver of the vehicle, a condition which reflects the current gradual expansion of the sharing economy.

Considering that an individual's utility of a certain mode will change according to the individual's socio-economic characteristics, the utility function formula was calculated and the model was estimated by differentiating the dummy variable for each mode. In

order to assess the goodness of fitness of the estimated model, ρ^2 values were compared. The results showed that the latent class model had the highest ρ^2 value of 0.2986 and the most explanatory power. To verify that the estimated model can be applied to other samples, the application process was performed using samples other than the ones used in the estimation. The results of the application showed that more than 50% of passengers actually use the most efficient mode. By comparing actual usage rates of certain modes with estimated usage rates, we confirmed that the latent class model showed the most similar usage rates.

In addition, to determine which model showed the most diversity in individual preferences, elasticities of demand according to increases in travel time, travel cost, and the value of travel time was obtained. The results show that the value of travel time changed according to the model used. In the mixed logit model, the value of travel time was larger in the truncated triangular distribution than in the truncated normal distribution. The triangular distribution had fewer extreme values than the normal distribution. As the values were spread evenly, relatively large values were more distributed, and the value of travel time appeared high. In the case of the latent class model, the value of travel time differed according to the characteristics of the individuals belonging to their class. Value of travel time would vary depending on what criteria were used to classify individuals.

As demand elasticities for travel time and travel costs were most sensitive in the latent class model, we know that the latent class model best reflects the diversity of individual preferences. If the

elasticities of travel time and travel cost for each mode are compared, the elasticity of the subway is found to be the least sensitive in all models, and in the majority of models and modes, time elasticity is greater than cost elasticity. Because this study targets commuters at morning peak hours and arriving at work on time is an important decision factor in their choice of mode, commuters are more sensitive to changes in time. Thus, even if the travel time or travel cost using the subway increases, the demand for a mode that is punctual decreases less than for other modes.

This study estimates a mode choice logit model that can reflect the diversity of individual preferences. Through analysis, the latent class model was found to have the most explanatory power, and this is relevant to the search for a model that can reflect the gradually increasingly segmental characteristics of individuals. Therefore, it is necessary to reform general mode choice logit models to ones that reflect the diversity of individual preferences.

Key Words: Mode Choice Model, Multinomial Logit Model, Mixed Logit Model, Latent Class Model, Preference Variation, Value of Travel Time

Student Number : 2016-24830